Restricted to VTI's trainees

(7)

Computer Hardware Repair & Maintenance

Trainee Manual-I

SUCCESS THROUGH SKILL
PUNJAB VOCATIONAL TRAINING COUNCIL

Punjab Vocational Training Council

Computer Hardware Repair & Maintenance

Trainee Manual

Punjab Vocational Training Council

FOREWORD

The Punjab Vocational Training Council (PVTC) was established in June 1998. The objective was to give need based demand driven technical training to enhance government noble endower of poverty alleviation. For this purpose PVTC has developed curricula for various courses to cope with the growing demand of industrial, agriculture and services sector, keeping in view the modern technology.

Due to IT revolution world has become global village. The promotion of applications of information technology is the trend in current world economic and social development and the key link to upgrading industrial structure and realizing industrialization and modernization.

The Punjab Vocational Training Council, as part of national policy intend to develop a strategy for IT infrastructure development, software development, value added services and capacity building to support his objectives of turning Pakistan into a knowledge society and also an active player in global economy through the use of technology. The emphasis is to make Pakistan a developed country in shortest possible time by using our human resource.

PVTC recognizing the increasing influence of IT in our country in the first instance, encouraged to acquire computer literacy as an initial step to build leadership, understanding and support for introduction and implementation of a comprehensive IT programme. The aim is to produce skilled worker to speed up the development of information industry and service industry, as promotion of information technology is the trend in current world economic and social development and key link to upgrading Pakistan's industrial and service structure. The information infrastructure is a preferential task for national economic and social development.

The information technology should be used extensively in all circles of society and use of computer and the Internet is wide spread. Information technology is being used in various sectors including governmental administration & management, public service, enterprise management and daily operations. The IT is being implemented in integration of networks of Telecommunication, TV, Banks, Shopping outlets and other business sectors.

The PVTC has taken the step to impart IT education and its application to business including industrial and commercial enterprise.

This curricula will go along way towards overall improvement of this standards of skill trainee and ultimately in poverty alleviation.

Thanks to the efforts of the following

1. Mr. Muhammad Shakil, Manager Curriculum, PVTC, Lahore.
2. Mr. Shahid Saleem, Senior Instructor, VTI, Township, Lahore.
3. Mr. Mujahid Ali Sheikh, Instructor, VTI, Green Town, Lahore.
4. Mr. Hassan Farooq, Instructor, VTI, Bahawal Nagar.

Suggestions for improving this curriculum are almost welcome.

SAJID NASEER KHAN
GENERAL MANAGER

SIKANDAR M. KHAN
CHAIRMAN PVTC

Table of Contents

Chapter 1

Introduction to Computers & Computer Maintenance Tools



OBJECTIVES

Introduction to Computers

The Development of Computers

Role of Computer Services Professional

What is Computer?

The Word Computer comes from the word "compute" which, means to calculate. So computer is normally considered to be a calculating device that can perform arithmetical operations at enormous speed. So we can say that like a calculator a computer and add, subtract, multiply and division etc. quickly and with high speed and accuracy.

Thus the computer can be defined as:

"It is an electronic device which accept instructions in the form of input and after processing it gives us information in the form of output."

Brief History of Computer

The History of Computer backs a very long way. In 17th Century BLAISE PASCAL a mathematician and scientist of France developed the first mechanical calculator, which he called the "Pascalins".

The designer of the first computer was CHARLES BABBAGE a mathematician of UK. Babbage in 1822 design a special type of calculator called difference Engine for the calculating the long list of number required in long table. In 1833 he developed a machine called "ANALYTICAL ENGINE".
Later on American Scientist Working at Howard University and developed a computer between 1937 and 1943 called Howard Mark-I.

First electronic computer was developed in 1946 at the University of Pennsylvania USA. And first Personal computer (PC) was developed in 1981. Now after a revolutionary development in PC Technology, Personal Computer affordable at and can solve our daily computing problems.

Classification of Computers

Computer is classified according to their size, Technology and the working principal into different categories.

1) Depending upon their size
2) Depending upon Working Principle
3) Depending upon Technology

1. Depending Upon Their Size:

Super Computer:

Fastest ,Most Powerful, Most expensive, Large in size and Memory, Multi –User.

Mainframe Computer:

Less powerful than Super Computer, Large in size and memory, Multi –user, Expensive.

Mini Computers:

Less powerful than Mainframe, Medium Sized, Multi-user, Expensive.

Microcomputer:

Single user, Powerful Enough, In expensive, Personal Computer

2. Depending upon their working Principal

(a) Analog Computer:

Such a type of computer in which input is given in the form of digits but the output is receiving in graphical form. This output is correct 90% to 95% (i.e. Wall Clock, Thermometer, Magnetic Needle etc.)

(b) Digital Computer

Digital computer that in which input is given in the form of digits and output is also received in form of digits (Digital Watch, Calculator etc.)

(c) Hybrid Computer

These computers have both properties analog and digital. Hybrid Computer given output in both digits and graphs all PC are the Hybrid computers having both properties.

3. Depending upon their Technology:

Computer has improved since those early days. Each major step forward is called a ‘Generation’. There are the following generations:

(1) First Generation (1945-1959):

The first generation of computers used electrically devices like vacuum tubes, capacitor. The machine was huge occupied a whole room and its processing speed was very slow. The ENIAC, ADVAC, EDSAC fall under this category.

(2) Second Generation (1959-1965):

These use transistor, which were faster and more reliable than vacuum tubes. Some of the computers in this category are UNIVAC-III and IBM-7070.

(3) Third Generation (1966-1971):

These computers were still smaller. They used integrated circuits (IC’s). It became known as the silicon chip or “chip”. Disk is used for storage purpose in this generation’s computers.

(4) Fourth Generation (1971-Present):

These computers were manufactured in 1970,s after further development in the chip. VLSIC (Very Large Scale integrated circuits) could be replaced by the Integrated circuit.

(5) Fifth Generation (In Future):

Fifth generation computer is still being developed. Research is being carried out to produce a computer with Artificial Intelligence. Such computers will be able to learn, adopt and correct mistakes. This development in hardware and software will lead to further advancement.

Introduction to Personal Computers

Personal computer is also called Microcomputer. Microcomputer was designed to be one person at a time. These are small in size and mainly used in Accounting, Database, Word Processing and Spread Sheet packages etc.

Intel Based microcomputer is of two categories;

Old Series: 8088, 8086, 80286, 80386, 80486 and 80586

New Series: New categories of PCs are:
Pentium I, Pentium II Pentium III, Celeron, AMD

PCs are available in branded and unbranded form.

Branded Computer:

These are marketed under a brand name. For example IBM, APPLE, ACER, COMPAQ, HEWLLETT PACKARD (HP), DELL and so on. These are high quality, high cost machines.

UN-Branded Computer:

These computers have no name and take the name of the seller. They are normally made in Taiwan or Korea. They are cheap and less reliable machines.

PC's normally come in four sizes & Shapes;

1. Desk Top PC
2. Tower PC
3. Laptop PC
4. Palm Top PC

Characteristics of a Computer

1 Speed:

A Computer is a very fast device. It can perform in a few seconds the amount of work that a human being can do in an entire year.

2 Accuracy:

The accuracy of a computer is consistently high and the degree of accuracy of a particular computer depends upon its design.

3 Diligence:

A computer is free from monotony, tiredness and lack concentrations. It can do work for hours.

4 Versatility:

Versatility is one of the most wonderful things about the computer. It can do many types of work such as preparing results, bills, letters etc.

5 No. IQ:

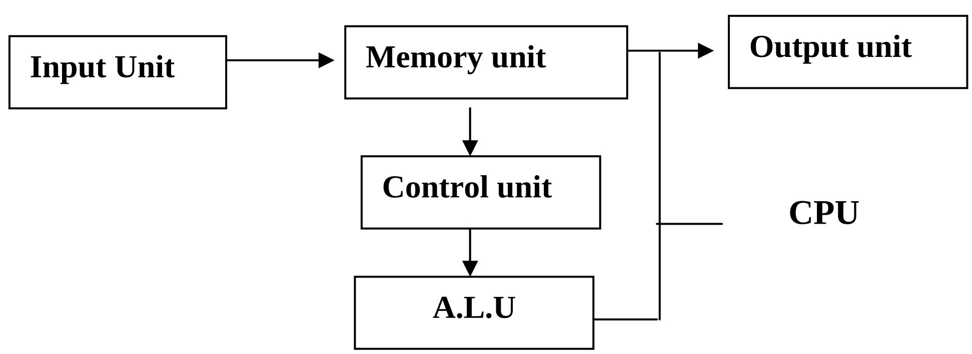
A computer cannot take its own decision in this regard.

6 No Feeling:

Computers have no feelings and no instinct because they are machines.

Components of Computer

Computer has following main components;



Input Unit:

Input Devices/ units are used to enter data and instructions into the computer. Before data or instructions are entered into computer these are converted into computer code and then entered them into computer memory. For example Keyboard, Mouse, Scanners etc.

Central Processing Unit:

Central processing unit is the brain of the computer. In a human body all major decisions are taken by brain and the other parts of the body function as directed by the brain. Similarly in the computer system all major calculations and comparisons are made inside the CPU and the CPU is also responsible for activating and controlling the operations of the other units of a computer. I t has following units

Output Devices:

The output devices are used to receive the results from the CPU in the form of electrical signals and convert them into readable form and their display the output. For example: Monitor, Printers, and Plotters Speakers (for sound).

Hardware:

Such physical equipments/components that make up the computer are called hardware (that we can see and touch). Computer hardware is the physical components of the computer. For Example: CPU, Monitor, Keyboard, Motherboard, Hard Disk & Mouse etc.

Software

Sets of programs that are used to communicate with computer are called software. Or the programs that we make in the computer are called software (that we can see but not touch). Languages and Packages. Like MS. Words, MS Excel, Basic & COBOL etc. Computer does nothing unless it is given instruction to perform.

Central Processing Unit:

CPU is the brain of computer system. All the computer processing takes place here. It is composed of three units.

1. Arithmetic and Logic Unit
2. Memory Unit
3. Control Unit

1. Arithmetic and Logic Unit:

The Arithmetic and Logic Unit performs arithmetic (addition, subtraction, multiplication and division) calculations and logical computation. ALU manipulates data.

2. Memory Unit

Data being process by ALU and intermediate result of processing or stored in main memory also called RAM. This memory is a temporary memory as soon as the electricity is switched off data in this memory is lost. This is volatile memory.

3. Control Unit

Control unit maintains order and controls processing activity in CPU. It directs the sequence of operations. It interprets instructions and produces signals for their execution. It also communicates with input and output devices. In personal computer the combination of ALU and control unit is called processor or microprocessor. The main memory is considered separately. So the CPU in a PC comprises of microprocessor and main memory.

How A CPU Works?

i. The control unit copies the programs and data into the main memory.
ii. Data and instructions are fed step by step into ALU by the control unit for processing.
iii. ALU process data. Control unit takes the result from ALU and puts them into main memory.
iv. From the main memory the results are moved for storage, display or printing as required.

Power of a Microprocessor

The power of a microprocessor mainly depends upon two parameters:

i. Speed
ii. Word Length

Speed of a Microprocessor:

Clock Speed:

The frequency of oscillations or clock cycles/sec. It is measured in mega hertz (MHz). PC’s today operate in the range of 100 MHZ to 1.7 GHZ.

Processing speed:

It is measured in million instructions per second (MIPS). Some instructions may take more than one cycle to operate.

Word Length

It is number of bits of data that can be processed by processor in a single step or cycle (For Example: processors can be 16 bits, 32 bits , 64 bits or 128 bits capacity).

Input Devices

The input devices that are used to give data to the computer and are called input devices. These devices are attached to the computer and controlled by CPU. These are also called peripheral devices. The mainly used input devices are:

- Keyboard.
- Mouse
- Scanner
- Digital Camera
- Joystick
- Microphone
- Light Pen

Keyboard::

It is the most familiar input device. It is like a typewriter. With alphabetic characters (A to Z) and numeric character (0 to 9). And some function keys (F1 to F12). F1 is normally used for help on a different subject. And other special characters are also on the key board.

Mouse:

Mouse is another input device used to point to various operations on the screen. It has two or three buttons on the top. At its bottom it has rubber or metal ball called magnet. The mouse is mainly used in the graphic applications it is also
used to play games on the computer.

Scanner

A scanner is also an input device. It is also called optical reader. It is capable of scanning or reading text printed on the paper, and directly enters into the computer's memory. This is a fast and accurate method of entering data into computer.

Digital Camera

It is very fast input device that captures the image and convert into computer memory. It is used for visual conference on the Internet.

Joystick

It is also an input device and is used for playing video games in computer. The games are easily controlled by joystick.

Microphone:

It is also called audio equipment. It is used to give the input in the form of sound or voice. It is used in voice talking.

Light Pen:

Light Pen is also an input device that is used for input in the form of drawing and designing. It is used in graphics packages for drawing purposes.

Output Devices

- Monitor
- Printers
- Plotters
- Speaker

Monitor

To display the result from a computer. A TV like device is called monitor. It is also called Video Display Unit (VDU) or it displays text and graphic objects.

Monitors are basically of two types.

1. Monochrome Monitor
2. Color Monitor

Monochrome Monitors

These Monitors are cheap and display only one color.

Color Monitors

These monitors display many colors. They are expensive than monochrome. The resolution of these monitors is measured in Pixels (Picture Elements). More the pixels better is the image displayed on the Screen. .

Printers

We can see output on a screen or can print out on a paper through a device called printers. Most of the printers produce output in black and white but some can print in colors. Epson, Hewlett Packard is popular printers names. Basically the printers are of these Types:

- Character Printers
- Line Printers
- Page Printers

Character Printers

Character Printers print a one character at a time. Examples of these are Dot Matrix Printers, and the Daisy Wheel Printers etc.

Line Printers

Line printers print one line at time. They are faster than character printers, and their printing speed ranges from 200 to 300 lines per minutes. They are used in offices where hard copy output INS large as for example case of offices bills (Gas, Electricity, Telephone Bill or even bank statements). Examples: Chain Printers, Drum Printers.

Page Printers

Page printers process the output one page at a time. Examples of these are Laser printer etc. These are the fastest printers available. They work on the same technique as a photocopier. They are relatively expensive but their output is of high quality.

Plotters

Plotters are used in computer Aided Design (CAD) packages. They are used to print high quality graphics. A Plotter produces drawing, maps, design and pictures. These are widely used in science, engineering and industrial designing.

Speakers

It is also an output device. It gives us the result or output in the form of sound. Specially used with audio software.



Storage Devices

During the processing a computer stored data in main memory. For permanent storage it needs auxiliary storage devices also called Secondary storage. Secondary storage can store large volumes of data. The unit normally used for secondary storage in Mega Byte (MB) and Gega byte (GB). These are mainly of three types:

- Magnetic Tape
- CD (Compact Disk)
- Magnetic Disk

Magnetic Tape

It is a continuous plastic strip wound on a reel, quite familiar to the tape used in reel-to-reel stereo recorders. The magnetic tap is plastic base is treated with a magnetizable coating. Typically the tape is on half inch in width. They are used as back up storage devices.

Magnetic tape has two parts:

- Recording Tape.
- Tape Drive.

Recording Type

It is made of plastic having magnetic material on it on which data is permanently stored. Data in stored in form of binary code (in the form of 0-1). On tape there are different tracks/frames and each track has different sectors.

Tape Drive

With the help of tape drive computer reads data from tape. It has two read a write heads. Read head reads the data and write head used to write that data in computer memory.

Disk



Disk is made of plastic or metal. It is also a permanent type of storage device. To store the data on a disk by a specific command disk is prepared for storage. Different tracks are made on disk. Each track is divided into further divisions called sectors and data is stored on both side of the disk. Each sector has a capacity of 512

bytes data storage.

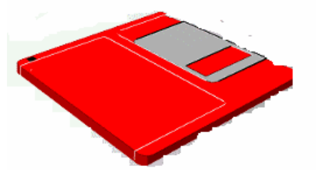
Types of Disk

i) Floppy Disk
ii) Hard Disk
iii) CD-ROM

Floppy Disk

It is made of plastic having magnetic material on it. It is enclosed in a plastic jacket. On the lower right of jacket there is a hole called notch. If we close the notch we cannot write on this disk but only can be read.

Classification of this disk is such as:



Size & Capacity of a Floppy Disk

Size	Density	Total	Sector Per Track	Total Capacity
5.25"	Double	80	8,9	360KB
5.25"	High	160	15	1.2MB
3.5"	Double	160	9	720KB
3.5"	High	160	18	1.44MB

Hard Disk

It is also called fixed disk. It consists of more them one metal plates fixed with spindle. Metal plates have magnetic material on them. Data is stored on both sides of plates. Its capacity is from MB to GB.



CD-ROM:

It is abbreviated from Compact Disk Read Only Memory. It is an optical laser disk that uses laser beams to store and retrieve data. Once the data has been written on CD cannot be changed or erased. Therefore it is called Read only Memory. It has high capacity storage device that can space more that 550-MB. It is most useful for storing reference data. For example we can find individual CDs of Encyclopedia, Oxford English Dictionary, Work of Shakespeare and so on.

What is Data?

A Collection of facts and figures that we give to the computer for gathering useful results and information is called data. There are three types of data.

- Alphabetic Data
- Numeric Data
- Alphanumeric Data

Alphabetic Data

This type of data consists of alphabets from (A to Z) or (a to z). This type of data is used to enter names. For example: Pakistan

Numeric Data

This data type consists of (0 to 9). For example (512, 586). This data type is used for mathematical calculations etc.

Alphanumeric Data

This type of data is a combination of Numeric and Alphabetic both. It is called character data. Example: 512-D is called character data type.

Units of data

Followings are units of the Data:

BIT:

The smallest unit of data is called bit. It is the abbreviation of Binary Integers. These are two digits '0' and '1'. The '1' states for 'ON' and '0' states for 'OFF'.

Nibble:

A combination of 4 bits is called one Nibble.
4bits= 1 Nibble.

Byte:

The bits in memory are arranged into a cell/byte. Byte is a combination of 8 bits. It is a small unit and also called one character.

8 bits = 1 Byte

Word:

A combination of more than one bytes is called word. Two bytes is called half word and 4 bytes is equal to full word. We can say that combination of characters is called the word.

Scale for measuring of data Units:

8 Bits	=	1 Byte
1024 bytes	=	1 Kilo Byte
1024 Kilo Byte	=	1 Mega Byte
1024 Mega Byte	=	1 Gega Byte
1024 Gega Byte	=	1 Terabyte

Software

Sets of programs that are used to communicate with computer are called software. Or the programs that we make in the computer are called software (that we can see but not touch). Languages and Packages. Like MS. Words, MS Excel, Basic & COBOL etc. Computer does nothing unless it is given instruction to perform.

Software has two types:

1) System Software
2) Application Software

1. System Software

System Software consists of the programs that control over all operations of the computer system.

System Software has following types:

- Operating System
- Translators
- System Utilities

Operating System:

A Software that controls the hardware 'OR' A Software that provides such an environment to user in which he operates computer easily. It is responsible for the flow of data between different components of computers. It has of two types:

i. Single User Operating System.
ii. Multi User Operating System.

Single User Operating System:

Such an operating system in which one user can work at a time. For Example: DOS, Windows 95, Windows 97, and Windows 98. etc.

Multi User Operating System:

Such an operating system on which more than one user can work called M. U. O. S. these are also called NOS For example: Novel, Unix, Linux, and XENIX, Window NT, etc.

Translators:

To execute programs from high level language to computer code. (The binary pattern '0' and '1'), some system programs are used called translators.

These are of following types:

i. Interpreters
ii. Compilers
iii. Assemblers

Interpreters:

Interpreters convert or translate the program one instruction at a time. Interpreter translates the program each time it is run. This makes it slower.

Compilers:

Compilers are also used to translate the high level language into computer code. They translate the whole program at a time. Unlike interpreters they make the execution of the program very fast.

Assemblers:

Assemblers convert the low level assembly language program into machine code.

System Utilities:

Utilities are system softwares. They provide useful services by performing common tasks such as sorting, merging and copying.

2. Application Software

Software designed to do a specific task is called application software.

It has two categories:

1) General Software
2) Standard Software

General Software:

It consists of different languages.

Language:

All the computers programs and instructions are written in computer language. It is such software in which the programmer makes the programming.

Types of language:

❖ Low level language

- High level language
- Middle level language

Low level language:

Such a language, which is comparatively close to the machine language or computer, code. Machine language.

High level language:

Nearly closet to human language like an English oriented language. Which does the computer not directly understand.

1. FORTRAN
2. COBOL
3. BASIC
4. PASCAL
5. PL/1
6. RPG
7. ALGOL
8. PROLOG

Middle level language:

Such a language which has both properties. In this language program is written in both low level and high level code. For example C, C++, languages

Standard Software :

Package:

Package is such a software in which programming is already made. User just operates this software. In other words we can say that package is a pre-made program.

It has following categories:

1) Word Processing
2) Spread Sheet
3) Database
4) Graphics

Word Processing:

Word processing software package can help you write and edit the document that you need. You can also retrieve store, safe, and preview and print it whenever needed. For example Microsoft Word, WordStar, WordPerfect, word pad, Notepad etc.

Spread Sheet:

Spreadsheets are software package in which data is written in the form of Rows & Columns. You can prepare different sheets (i.e. salary sheet, result sheet) and can use different formulas and can also draw a chart E.g. Excel, Lotus etc.

Database:

Collection of related data files. A data file consists of information stored in the term of records and fields. A record is a set of logical related field. A field is one item in a record. For example MS-Access, Oracle, Dbase etc.

Graphics:

Graphics packages getting more popularity with the passage of time. They require a lot of memory and high-resolution monitors who are capable of displaying a large number of dots per inch on the screen. There are many graphic packages like as Auto CAD, Corel Draw, Instant Artist and Paintbrush.

Device Driver Software

PC has different devices attached to it like as input, output, storage and some other components. Each device has a specific software which is installed to drive or operate it is called device driver software. Because different companies designed same equipment and each company’s device has different software. The manufacturers, on a floppy disk or a CD-ROM provide such software. It is essential to copy or install such software onto computer hard disk before the device can work properly.

Firm Ware:

A set of software or similar instructions forming a more or less permanent and unerasable part of a computer's memory.

Share Ware:

Software readily available for a nominal fee and provides communication between hardware & software.

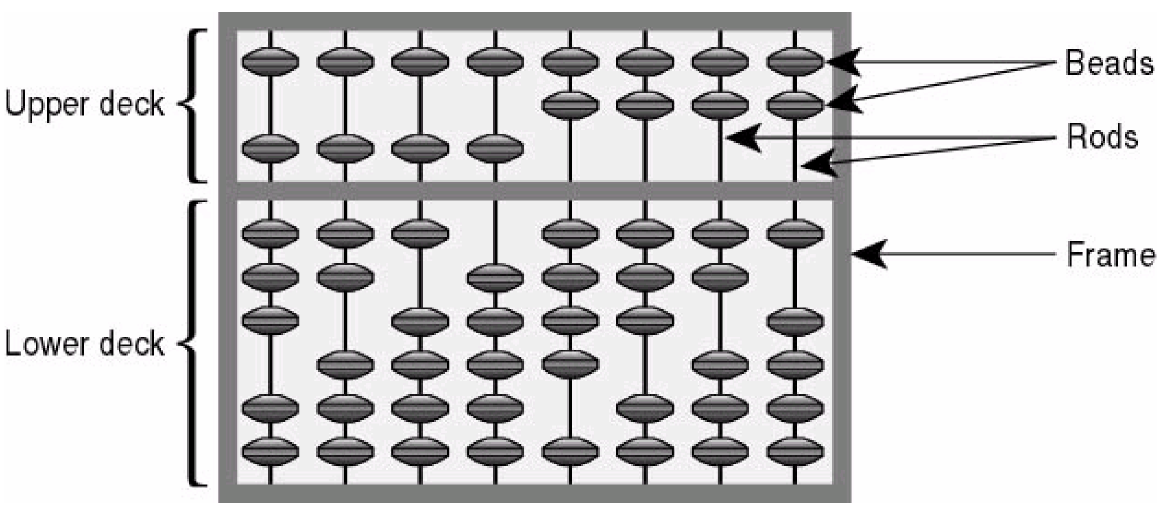
The Development of the Computer

In this lesson, we take a brief look at the development of the computer. By understanding its origins, you'll gain an appreciation for both the complexity and simplicity of today's computers.

Many of us think only in terms of electronic computers, powered by electricity. (If you can't plug it in, is it a computer?) But as the definition in *Funk & Wagnalls Standard College Dictionary* makes clear, to "compute" is to "ascertain (an amount or number) by calculation or reckoning." In fact, the first computers were invented by the Chinese about 2500 years ago. They are called *abacuses* and are still used throughout Asia today.

The Abacus

The abacus, shown in Figure 1.1, is a calculator; its first recorded use was circa 500 B.C. The Chinese used it to add, subtract, multiply, and divide. But the abacus was not unique to the continent of Asia; archeological excavations have revealed an Aztec abacus in use around 900 or 1000 A.D.



The first computer

The Analytical Engine (A Pre-Electronic Computer)

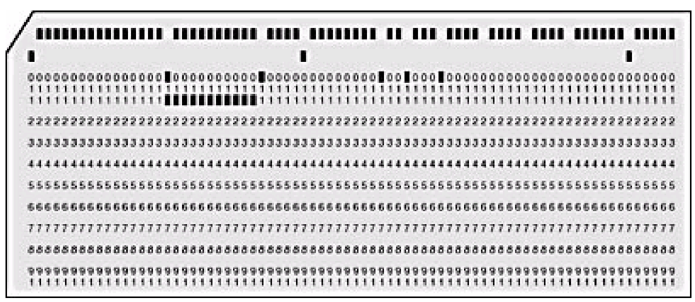
The first mechanical computer was the analytical engine, conceived and partially constructed by Charles Babbage in London, England, between 1822 and 1871. It was designed to receive instructions from punched cards, make calculations with the aid of a memory bank, and print out solutions to math problems. The precise work needed to engineer its thousands of moving parts was beyond the ability of the technology of the

day to carry out. It is doubtful whether Babbage's brilliant concept could have been realized using the available resources of his own century. But if it had been, it seems likely that the analytical engine could have performed the same functions as many early electronic computers.

The First Electrically Driven Computer

The first computer designed expressly for data processing was patented on January 8, 1889, by Dr. Herman Hollerith of New York. The prototype model of this electrically operated tabulator was built for the U. S. Census Bureau and computed results in the 1890 census.

Using punched cards containing information submitted by respondents to the census questionnaire, the Hollerith machine made instant tabulations from electrical impulses actuated by each hole. It then printed out the processed data on tape. Dr. Hollerith left the Census Bureau in 1896 to establish the Tabulating Machine Company to manufacture and sell his equipment. The company eventually became IBM, and the 80-column punched card used by the company, shown in Figure 1.2, is still known as the Hollerith card.



Typical 80-character punched card

The Digital Electronic Computer

The first modern digital computer, the ABC (Atanasoff-Berry Computer), was built in a basement on the Iowa State University campus in Ames, Iowa, between 1939 and 1942. This machine included many features still in use today: binary arithmetic, parallel processing, regenerative memory, separate memory, and computer functions.

The first large-scale digital computer, the Electronic Numerical Integrator and Computer (ENIAC). It was built at the University of Pennsylvania's Moore School of Electrical Engineering. Begun as a classified military project, ENIAC was destined to prepare firing and bombing tables for the U.S. Army and Navy. Finally assembled in 1945.

Although programming ENIAC was a mammoth task requiring manual switches and cable connections, it became the workhorse for the solution of scientific problems from 1949 to 1952. ENIAC is considered the prototype for most of today's computers.

Another device important to computer history is the COLOSSUS I, an early digital computer built at a secret government research establishment at Bletchley Park, Buckinghamshire, England. COLOSSUS I was designed for a single purpose: BREAKING CODES. Using punched paper tape input, it scanned and analyzed 5000 characters per second. Colossus became operational in December 1943 and proved to be one of the most important technological aids to victory in World War II. It enabled the British to break the otherwise impenetrable German "Enigma" codes.

The 1960s and 1970s marked the era of the MAINFRAME COMPUTER. The following table provides an overview of the evolution of modern computers—it is a timeline of important events.

Year	Events
1971	The 4004—the first 4-bit microprocessor—is introduced by Intel. It boasts 2000 transistors with a clock speed of up to 1 MHz (megahertz).
1972	The first 8-bit microprocessor—the 8008—is released.
1974	The 8080 microprocessor is developed. This improved version of the 8008 becomes the standard from which future processors will be designed.
1975	Digital Research introduces CP/M—an operating system for the 8080. The combination of software and hardware becomes the basis for the standard computer.
1976	Zilog introduces the Z80—a low-cost microprocessor (equivalent to the 8080). The Apple I comes into existence, although it is not yet in wide use.
1977	The Apple II and the Commodore PET computers, both of which use a 6502 processor, are introduced. These two products become the basis for the home computer. Apple's popularity begins to grow.
1978	Intel introduces a 16-bit processor, the 8086, and a companion math coprocessor, the 8087. Intel also introduces the 8088. It is similar to the 8086, but it transmits 8 bits at a time.

Year	Events
1980	Motorola introduces the 68000—a 16-bit processor important to the development of Apple and Atari computers. Motorola's 68000 becomes the processor of choice for Apple.
1981	The IBM personal computer (PC) is born; it contains a 4.7-MHz 8088 processor, 64 KB (kilobytes) of RAM (random access memory) and is equipped with a version of MS-DOS 1.0 (three files and some utilities). Available mass-storage devices include a 5.25-inch floppy drive and a cassette tape drive.
1982	Intel completes development of the 80286—a 16-bit processor with 150,000 transistors. MS-DOS 1.1 now supports double-sided floppy disks that hold 360 KB of data.
1983	IBM introduces the XT computer with a 10-MB hard-disk drive. MS-DOS 2.0 arrives—it features a tree-like structure and native support for hard-disk drive operations.
1984	The first computer with an 80286 chip—the IBM AT—enters the market. It is a 6-MHz machine with a 20-MB hard-disk drive and a high-density, 1.2-MB 5.25-inch floppy-disk drive.
1985	MS-DOS 3.2, which supports networks, is released.
1986	The first Intel 80386-based computer is introduced by Compaq; it features a 32-bit processor with expanded multitasking capability (even though no PC operating system yet fully supports the feature).
1987	MS-DOS 3.3 arrives, allowing use of 1.44-MB 3.5-inch floppy-disk drives and hard-disk drives larger than 32 MB.
1988	IBM introduces the PS/2 computer series. A complete departure from previous machines, its proprietary design does not support the hardware and software available on IBM PCs or clones. Microsoft (with the help of IBM) develops OS/2 (Operating System 2), which allows 32-bit operations, genuine multitasking, and full MS-DOS compatibility. Microsoft releases MS-DOS 4.0.

Year	Events
1989	Intel introduces the 80486 processor; it contains an on-board math coprocessor and an internal cache controller (offering 2.5 times the performance of a 386 processor with a supporting coprocessor).
1991	MS-DOS 5.0 offers a significantly improved DOS shell.
1992	The Intel i586 processor, the first Pentium, is introduced, offering 2.5 times the performance of a 486. IBM expands OS/2, and Microsoft Windows is introduced.
1993	MS-DOS 6.0 arrives. The term “multimedia” (the inclusion of CD-ROM drives, sound cards, speakers, and so forth, as standard equipment on new personal computers) comes into use.
1994	Intel delivers the first 100-MHz processor. COMPAQ Computer Corporation becomes the largest producer of computers.
1995	Windows 95, code-named Chicago, is introduced by Microsoft. It features 32-bit architecture. IBM has now shipped over one million OS/2 Warp software packages. The Internet, having expanded far beyond its beginnings as a network serving government and university institutions, is now in everyday use by the rapidly growing population with access to a modem. Computer prices drop as performance increases. IBM purchases Lotus (maker of the popular Lotus1-2-3 spreadsheet).
1995–1996	Software manufacturers scramble to make their products compatible with Windows 95.
1997	Microprocessor speeds exceed the 200-MHz mark. Hard-disk drive and memory prices fall while basic system configuration sizes continue to increase. CD-ROM drives and Internet connections have become standard equipment for computers.
1998-2000	Personal computer performance continues to soar, and PC prices continue to fall. CPU speeds exceed 450-700 MHz, and motherboard bus speeds reach 100-166 MHz.

Year	Events
	Multimedia and Internet connections have become the de facto standard for new PCs. Universal Serial Bus (USB) is introduced. Windows 98 becomes the standard operating system for most new personal computers.
2001-2003	Processors introduced by Intel & AMD ranging 1 GHz---2.2 GHz. With 400-500 motherboard bus speed. Microsoft Launched Windows XP 2002 with variety of different features, Linux has been getting popular among expert users.

The Role of a Computer Service Professional

Matching the rapid pace of change in the industry, the role of the computer professional is constantly changing, too. Not too many years ago, all that was needed to repair a computer was a screwdriver, needle-nose pliers, the documentation for the computer, a boot disk with a few utilities, and a good MS-DOS reference manual. Although these tools are still fundamental to the job, to be ready for all situations, a technician has to travel with an entire library of technical manuals (a laptop computer with a CD-ROM drive and modem helps, too). The screwdriver is still the standard repair tool, but the technician is confronted with a wider array of case types, motherboard designs, processor types, and operating systems and a wider array of customer needs. Today's computer professional needs to be a technician, a scholar, and a diplomat rolled into one, as you can see by the table that follows.

Title	Skills
Technician	You are able to troubleshoot and repair hardware and software efficiently and quickly.
Scholar	You have the wisdom and perseverance to seek answers to what you don't know and build your base of knowledge. Learning never stops.
Diplomat	You are able to instill in the user (your customer) the confidence that you are in control and can fix things, even when you are encountering them for the first time. You are able to resolve the problem, even if your customer's (lack of) understanding of the computer might be part of that problem.

Computer Maintenance Tools

This Unit describes following different types of tools for troubleshooting and maintaining the Computer and Network systems: -

- Hand tools.
- Small nylon wire ties for "dressing" or organizing wires.
- Foam swabs, or lint-free cotton swabs
- Soldering iron and solder sucker.
- ESD (Electrostatic Discharge) Protection
- Test Equipment
    - Digital Multi-meter
    - Loop back adaptors
    - Logic Probe and Logic Pulser
    - POST Probe Card
- Networking
    - Crimping Tool
    - Punch down tool
    - Network Cable Tester
- Chemicals
    - Contact Cleaners
    - Component Freeze Sprays
    - Silicone Lubricants
- Vacuum Cleaners
- Diagnostics software

Hand Tools

The following list mentions some specific tools anyone should have if they intend to get inside their PC:

6mm nut driver
8mm nut driver
Small Phillips screwdriver
Small straight edge screwdriver
Medium Phillips screwdriver
Medium straight edge screwdriver

Chip extractor

Larger processors require a chip extractor if they are mounted in the older low insertion force (LIF) socket.

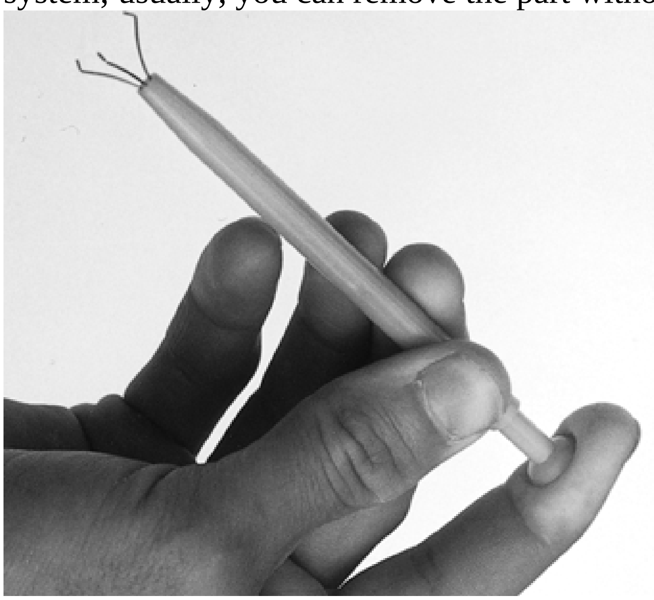
Chip inserter

Tweezers

The tweezers and parts grabber can be used to hold any small screws or jumper blocks that are difficult to hold in your hand.

Claw-type parts grabber

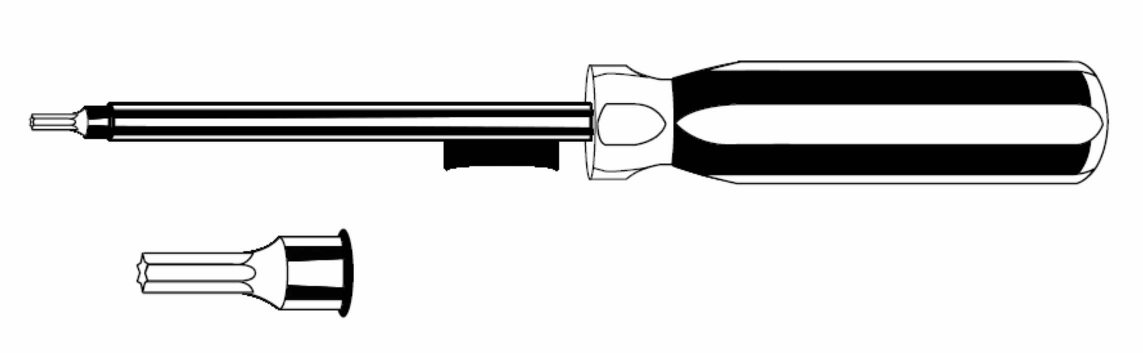
The parts grabber is especially useful when you drop a small part into the interior of a system; usually, you can remove the part without completely disassembling the system.



Parts Grabber

T10 and T15 Torx drivers

The Torx driver is a star-shaped driver that matches the special screws found in most Compaq systems and in many other systems as well.



Torx Driver

Needle-nose pliers

Pliers are useful for straightening pins on chips, applying or removing jumpers, crimping cables, or grabbing small parts.

Hemostat

Hemostats are especially useful for grabbing small components, such as jumpers.

File

You can use the file to smooth rough metal edges on cases and chassis and to trim the faceplates on disk drives for a perfect fit.

Wire cutter or wire stripper

The wire cutter or stripper, obviously, is useful for making or repairing cables or wiring.

Small flashlight

You might often need a flashlight to illuminate system interiors, especially when the system is cramped and the room lighting is poor, or when you are working on a system underneath a user’s desk.

ESD Protection Tools

Another consideration for computer toolkit is an electrostatic discharge (ESD) rotection kit. This kit consists of a wrist strap with a ground wire and an especially conductive mat with its own ground wire. Using a kit like this when working on a system helps ensure that you never accidentally damage any of the components with a static discharge. You

also can get just the wrist strap or the antistatic mat separately. A wrist strap is shown in the picture.

Note
It is possible (but not recommended) to work without an ESD protection kit if you're disciplined and careful about working on systems. If you don't have an ESD kit available, you should discharge yourself by touching some metallic part of the case while the computer is still plugged into the AC power source; then unplug the computer. Many systems today continue to feed power to the motherboard through the soft off connection whenever the computer is plugged in, even when the power switch is turned off. It can be very dangerous to work inside a PC that is still connected to a power source.



A typical ESD wrist strap clipped to a non-painted surface in the case chassis.

Soldering and De-soldering Tools

Soldering irons are used to make some necessary repairs though most of the maintenance work involves replacing the entire failed components. Simple soldering skills can save a lot of money in these cases. Majority of modern motherboards include I/O components, such as serial and parallel ports. Many of these ports are fuse-protected on the board which is normally a small soldered-in component. These fuses are designed to protect the motherboard circuits from being damaged by an external source. If a short circuit or static charge from an external device blows these fuses, the motherboard can be saved if these soldered-in fuses can be replaced. A low-wattage soldering iron of about 25 watts may be used perform minor repairs like above. More than 30 watts generates too much heat and can damage the components on the board. The soldering must be done quickly and a heat sink clipped to the component being soldered be used to save it from excess heat build-

up. A *heat sink* is a small metal clip-on device designed to absorb excessive heat before it reaches the component the heat sink is protecting.
In some cases, you can use a pair of hemostats as an effective heat sink when you solder a component. A solder sucker with a soldering iron can be used to remove components soldered into place on printed circuit board. When a device is required to be removed from a board, a soldering iron is used to heat the underside of the board and the point at which one of the component leads joins the circuit board until the solder melts. As soon as melting occurs solder-sucker nozzle is moved into the position and the actuator is pressed. When the plunger retracts, it creates a momentary suction that draws the liquid solder away from the connection and leaves the component lead dry in the hole. Always perform the heating and suctioning from the underside of a board, not from the component side. Repeat this action for every component lead joined to the circuit board.

Test Equipment

The following test equipment adds to the abilities of the hand tools used for computer maintenance.

Digital Multi-meter

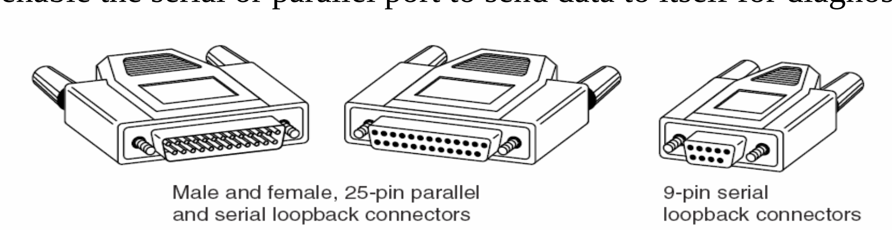
A multi-meter can be used for checking voltage signals at different points in a system, testing the output of the power supply, and checking for continuity in a circuit or cable. The DMM has a pair of wires called test leads or probes. The test leads make the connections so that you can take readings. Depending on the meter's setting, the probes measure electrical resistance, direct-current (DC) voltage, or alternating-current (AC) voltage. Each system-unit measurement setting has several ranges of operation. For example, DC voltage can be read in several scales, to a maximum of 200 millivolts (mv), 2v, 20v, 200v, and 1,000v. Since computers use both +5v and +12v for various operations, the 20v maximum scale should be used for voltage measurements. Making these measurements on the 200mv or 2v scale could possibly damage it because the voltage would be much higher than expected. Using the 200v or 1,000v scale works, but the readings at 5v and 12v are so small in proportion to the maximum that the accuracy is low. If the actual voltages are no known, start at the highest scale and then gradually go down. Most of the better meters have autoranging capability—the meter automatically selects the best range for any measurement. The meter selects the correct voltage range and displays the value.

Warning
Whenever you are using a multimeter to test any voltage that could potentially be 220V or above, always use one hand to do the testing, not two. Either clip one lead to one of the sources and probe with the other or hold both leads in one hand. If you hold a lead in each hand and accidentally slip, you can very easily become a circuit, allowing power to conduct or flow through you. When power flows from arm to arm, the path of the current is directly across the heart. Hearts tend to quit working when subjected to high voltages. .

Warning
You should be aware that many analog meters can be dangerous to digital circuits. These meters use a 9v battery to power the meter for resistance measurements. If you use this type of meter to measure resistance on some digital circuits, you can damage the electronics because you essentially are injecting 9v into the circuit. The digital meters universally run on 3v–5v or less.

Loopback Connectors (Wrap Plugs)

Loopback connectors are required for diagnosing serial- and parallel-port problems, you need (also called wrap plugs), which are used to circulate, or wrap, signals. The plugs enable the serial or parallel port to send data to itself for diagnostic purposes.



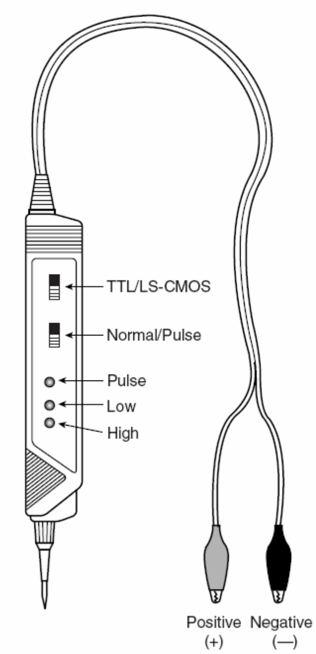
Typical wrap plugs including 25-pin and 9-pin serial and 25-pin parallel versions.

Various types of loopback connectors are available. To accommodate all the ports you might encounter, you need one for the 25-pin serial port, one for the 9-pin serial port, and one for the 25-pin parallel port.

Logic Probes and Logic Pulsers

A logic probe can be useful for diagnosing problems in digital circuits. In a digital circuit, a signal is represented as either high (+5v) or low (0v). Because these signals are present for only a short time (measured in millionths of a second) or oscillate (switch on and off) rapidly, a simple voltmeter is useless. A logic probe is designed to display these signal conditions easily.

Logic probes are especially useful for troubleshooting a dead system. By using the probe, you can determine whether the basic clock circuitry is operating and whether other signals necessary for system operation are present. In some cases, a probe can help you cross-check the signals at each pin on an integrated circuit chip. You can compare the signals present at each pin with the signals a known good chip of the same type would show—a comparison that is helpful in isolating a failed component.



A typical logic probe.

Logic probes also can be useful for troubleshooting some disk drive problems by enabling you to test the signals present on the interface cable or drive-logic board. A companion tool to the probe is the logic pulser. A *pulser* is designed to test circuit reaction by delivering a logical high (+5v) pulse into a circuit, usually lasting from 1 1/2 to 10 millionths of a second. Compare the reaction with that of a known-functional circuit. This type of device normally is used much less frequently than a logic probe, but in some cases it can be helpful for testing a circuit.

Chemicals

Chemicals can be used to help clean, troubleshoot, and even repair a system. You can use several types of cleaning solutions with computers and electronic assemblies. Most fall into the following categories:

Standard cleaners
Contact cleaner/lubricants

Standard Cleaners

For the most basic function—cleaning components, electrical connectors, and contacts—Alternative cleaning solutions are available in a variety of types and configurations. You can use pure isopropyl alcohol, acetone, Freon, trichloroethane, or a variety of other chemicals. You must exercise care when using organic cleaners, however, because they can cause swelling of some plastics, especially silicone rubber and PVC. You should make sure your cleaning solution is designed to clean computers or electronic assemblies.

Contact Cleaner/Lubricants

These chemicals are similar to the standard cleaners but include a lubricating component. The lubricant eases the force required when plugging and unplugging cables and connectors, reducing strain on the devices. The lubricant coating also acts as a conductive protectant that insulates the contacts from corrosion. These chemicals can greatly prolong the life of a system by preventing intermittent contacts in the future.

Silicone Lubricants

You can use a silicone lubricant such as WD-40 to lubricate the door mechanisms on floppy disk drives and any other part of the system that might require clean, non-oily lubrication. Other items you can lubricate are the disk-drive–head slider rails and even printer-head slider rails, to provide smoother operation. Using silicone instead of conventional oils is important because silicone does not gum up and collect dust and other debris. Always use the silicone sparingly. Do not spray it anywhere near the equipment because it tends to migrate and will end up where it doesn’t belong (such as on drive heads). Instead, apply a small amount to a toothpick or foam swab and dab the silicone lubricant on the components where needed. You can use a lint-free cleaning stick soaked in silicone to lubricate the metal print head rails in a printer.

Vacuum Cleaners

It is preferable to use a vacuum cleaner instead of air blower for cleaning a system. A vacuum cleaner is more useful when you are cleaning a system loaded with dust and dirt. You can use the vacuum cleaner to suck out the dust and debris instead of simply blowing it around on the other components. Tiny vacuum cleaners also are available for system cleaning. Some special vacuum cleaners are specifically designed for use on and around electronic components; they are designed to minimize electrostatic discharge (ESD) while in use. If you are using a regular vacuum cleaner and not one specifically designed with ESD protection, you should take precautions, such as wearing a grounding wrist strap. Also, if the cleaner has a metal nozzle, be careful not to touch it to the circuit boards or components you are cleaning.

Brushes and Swabs

You can use a small makeup, photographic, or paint brush to carefully loosen the accumulated dirt and dust inside a PC before using air blower or using the vacuum cleaner. Be careful about generating static electricity. In most cases, you should not use a brush directly on circuit boards, but only on the case interior and other parts, such as fan blades, air vents, and keyboards. Wear a grounded wrist strap if you are brushing on or near any circuit boards, and brush slowly and lightly to prevent static discharges from occurring. Use cleaning swabs to wipe off electrical contacts and connectors, disk drive heads, and other sensitive areas. The swabs should be made of foam or synthetic chamois material that does not leave lint or dust residue. Unfortunately, proper foam or chamois cleaning swabs are more expensive than typical cotton swabs. Do not use cotton swabs because they leave cotton fibers on everything they touch. Cotton fibers are conductive in some situations and can remain on drive heads, which can scratch disks.

Warning

One item to avoid is an eraser for cleaning contacts. Many people have recommended using a soft pencil type eraser for cleaning circuit-board contacts. Testing has proven this to be bad advice for several reasons. One reason is that any such abrasive wiping on electrical contacts generates friction and an ESD. This ESD can be damaging to boards and components, especially the newer low-voltage devices. These devices are especially static sensitive, and cleaning the contacts without a proper liquid solution is not recommended. Also, the eraser will wear off the gold coating on many contacts, exposing the tin contact underneath, which rapidly corrodes when exposed to air.

Some companies sell pre-moistened contact cleaning pads soaked in a proper contact cleaner and lubricant. These pads are safe to wipe on conductors and contacts with no likelihood of ESD damage or abrasion of the gold plating.

Chapter 2

Understanding Electronics Communication

OBJECTIVES

Computer Communication
Computer Bus

Computer Communication

Communicating is the act of giving, transmitting, or exchanging information. A key element in developing a device such as a computer is establishing a method of communication, both internally (for the transfer of information between hardware components) and externally (with the outside world). In this chapter, we discuss how a computer processes data and communicates (transmits information) with its user. Understanding this process is fundamental to understanding how computers work.

Early Forms of Communication

Humans communicate primarily through words, spoken and written. From ancient times until about 150 years ago, messages were either verbal or written in form. Getting a message to a distant recipient was often slow, and sometimes the message (or the messenger) got lost in the process.

As time and technology progressed, people developed devices to communicate faster over greater distances. Items such as lanterns, mirrors, and flags were used to send messages quickly over an extended visual range.

Dots and Dashes, Bits and Bytes

Telegraphs and early radio communication used codes for transmissions. The most common, Morse code (named after its creator, Samuel F. B. Morse), is based on assigning a series of pulses to represent each letter of the alphabet. These pulses are sent over a wire in a series. The operator on the receiving end converts the code back into letters and words. Morse code remained in official use for messages at sea almost to the end of the twentieth century it was officially retired in late 1999.

Morse used a code in which any single transmitted value had two possible states: either a dot or a dash. By combining the dots and dashes into groups, an operator was able to represent letters, and by stringing them together, words. That form of on-off notation can also be used to provide two numbers, 0 and 1. Zero represents no signal, or off; and one represents a signal, or on, state.

This type of number language is called *binary notation* because it uses only two digits, usually 0 and 1. It was first used by the ancient Chinese, who used the terms *yin* (empty) and *yang* (full) to build complex philosophical models of how the universe works.

Our computers are complex switch boxes that have two states and use a binary scheme as well. The value of a given switch's state—on or off—represents a value that can be used

as a code. Modern computer technology uses terms other than yin and yang, but the same binary mathematics creates virtual worlds inside our modern machines.

The Binary Language of Computers

The binary math terms that follow are fundamental to understanding PC technology.

Bits

A *bit* is the smallest unit of information that is recognized by a computer: a single on/off event, and is the abbreviation of Binary digIT.

Bytes

A *byte* is a group of eight bits. A byte is required in order to represent one character of information. Pressing one key on a keyboard is equivalent to sending one byte of information to the CPU (the computer's central processing unit). A byte is the standard unit by which memory is measured in a computer—values are expressed in terms of kilobytes (KB) or megabytes (MB). The table that follows lists units of computer memory and their values.

Memory Unit	Value
Bit	Smallest unit of information, shorthand term for binary digit
Nibble	4 bits (Half of a byte)
Byte	8 bits (Equal to one character)
Word	16 bits on most personal computers (longer words possible on larger computers)
Kilobyte (KB)	1024 bytes
Megabyte (MB)	1,048,576 bytes (Approximately one million bytes or 1024 KB)
Gigabyte (GB)	1,073,741,824 bytes (Approximately one billion bytes or 1024 MB)
Terabyte (TB)	1024 Gigabyte of data

The Binary System

The binary system of numbers uses the base of 2 (0 and 1). As described earlier, a bit can exist in only two states, on or off. When bits are represented visually:

- 0 (zero) equals off.
- 1 (one) equals on.

The following is one *byte* of information in which all eight *bits* are set to zero. In the binary system, this sequence of eight zeros represents a single character—the number 0.

0	0	0	0	0	0	0	0

The binary system is one of several numerical systems that can be used for counting. It is similar to the decimal system, which we use to calculate everyday numbers and values. The prefix "dec" in the term "decimal system" comes from the Latin word for ten and denotes a base of ten. That is, the decimal system is based on the ten numbers zero through nine. The binary system has a base of two, the numbers zero and one.

Counting in Binary Notation

Every schoolchild learns to count using the decimal system. There, the rightmost whole number (the number to the left of the decimal point) is the "digits" column. Numbers written there have a value of zero to nine. The number to the left of the digits column (if present) is valued from ten to ninety—the "tens" column. Ten is the factor of each additional row in the decimal system of notation. To get the total value of a number, we add together all columns in both systems: 111 is the sum of 100+10+1.

In our more common decimal notation, the values of numbers are founded on a base of ten, starting with the rightmost column. Any number in that position can have a value ranging from zero to nine. In the next column to the left, the values range from 10 to 99; and in the column to the left of that, values range from 100 to 999. Binary notation uses a system of right-to-left columns of ascending values, but in which each row has only two-instead of 10-possible numbers.

Under the binary system, the first row to the right can be only zero or one; the next row to the left can be two or three (if a number exists in that position). The columns that follow have values of four, then eight, then sixteen, and so on, each column doubling the possible value of the one to its right. Two is the factor used in the binary system, and just like decimal—zero is a number counted in that tally. Examples of bytes of information (eight rows) follow.

Byte—Example A

The value of this byte is zero because all bits are off (0 = off).

```
0      0      0      0      0      0      0      0      8      bits
128    64     32     16     8      4      2      1      #      values
```

Byte—Example B

In this example, two of the bits are turned on (1 = on). The total value of this byte is determined by adding the values associated with the bit positions that are on. This byte represents the number 5 (4 + 1).

```
0      0      0      0      0      1      0      1      8      bits
128    64     32     16     8      4      2      1      #      values
```

Byte—Example C

In this example, two different bits are turned on to represent the number 9 (8 + 1).

```
0      0      0      0      1      0      0      1      8      bits
128    64     32     16     8      4      2      1      #      values
```

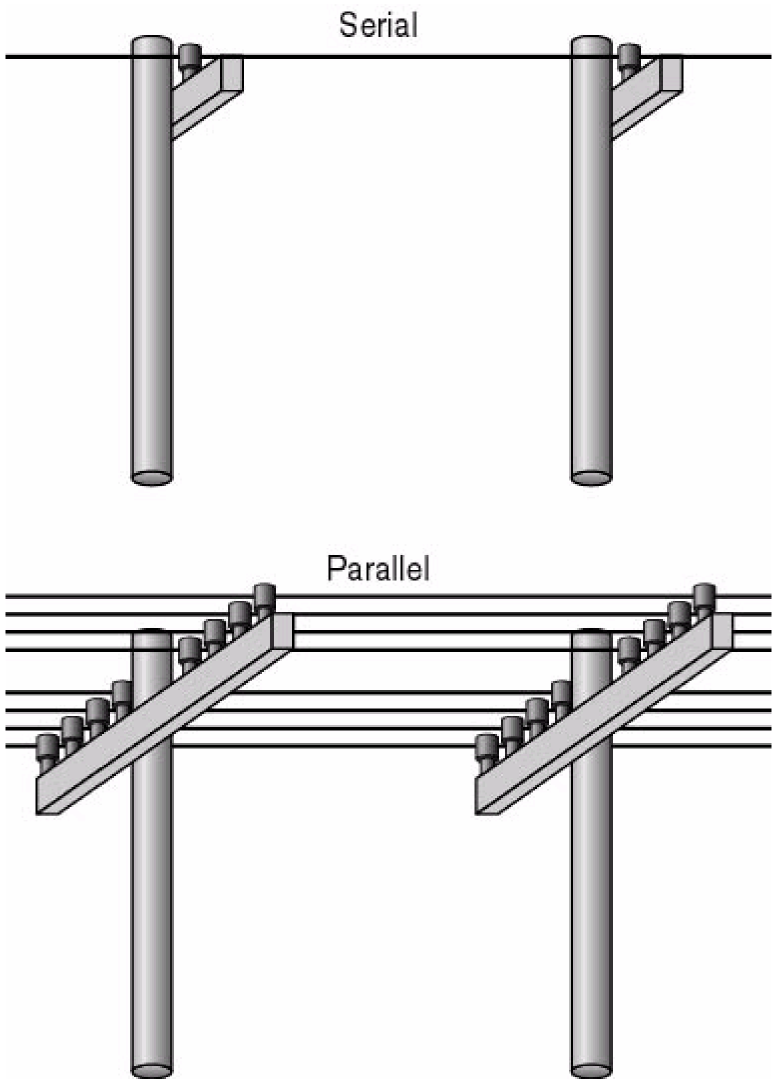
Because computers use binary numbers and humans use decimal numbers, hardware technicians must be able to perform simple conversions. The following table shows decimal numbers and their binary equivalents (0 to 9). You will need to know this information. The best way to prepare is to learn how to add in binary numbers, rather than merely memorizing the values.

Numbers are fine for calculating, but today's computers must handle text, sound, streaming video, images, and animation as well. To handle all of that, standard codes are needed to translate between binary machine language and the type of data being represented and presented to the human user. The first common code-based language was developed to handle text characters.

Parallel and Serial Devices

The telegraph and the individual wires in our PCs are serial devices. This means that only one element of code can be sent at a time. Like a tunnel, there is only room for one person to pass through at one time. All electronic communications are at some level serial, because a single wire can have only two states: on or off. To speed things up, we can add more wires. This allows simultaneous transmission of signals. Or, to continue our analogy, it's like adding another set of tunnels next to the first one; we still have only one person per tunnel, but we can get more people through because they are traveling in parallel. That is the difference between parallel and serial data transmission. In PC technology, we often string eight wires in a parallel set, allowing eight bits to be sent at once. This means that a single "send" can represent up to 256 numbers $2^8 = 256$. That is

the same number of values found in the ASCII code system. The picture below illustrates serial and parallel communication.



Serial and parallel communication

ASCII Code

The standard code for handling text characters on most modern computers is called *ASCII* (American Standard Code for Information Interchange). The basic ASCII standard consists of 128 codes representing the English alphabet, punctuation, and certain control characters. Most systems today recognize 256 codes: the original 128, plus an additional 128 codes called the *extended character set*.

Remember that a byte represents one character of information; four bytes are needed to represent a string of four characters. The following four bytes represent the text string 12AB (using ASCII code):

```
00110001    00110010    01000001    01000010
   1           2           A           B
```

The following illustrates how the binary language spells the word "binary":

```
    B          I          N           A           R           Y
01000010   01001001   01001110    01000001    01010010    01011001
```

It is very important to understand that in computer processing the "space" is a significant character. All items in a code must be set out for the machine to process. Like any other character, the space has a binary value that must be included in the data stream. In computing, the absence or presence of a space is critical and sometimes causes confusion or frustration among new users. Uppercase and lowercase letters also have different values. Some operating systems (for example, UNIX) distinguish between them for commands, while others (for example, MS-DOS) translate the uppercase and lowercase into the same word no matter how it is cased.

The following table is a complete representation of the ASCII character set. Even in present-day computing, laden with multimedia and sophisticated programming, ASCII retains an honored and important position.

Symbol	Binary 1 Byte	Decimal	Symbol	Binary 1 Byte	Decimal
0	00110000	48	V	01010110	86
1	00110001	49	W	01010111	87
2	00110010	50	X	01011000	88
3	00110011	51	Y	01011001	89
4	00110100	52	Z	01011010	90
5	00110101	53	A	01100001	97
6	00110110	54	B	01100010	98
7	00110111	55	C	01100011	99
8	00111000	56	D	01100100	100

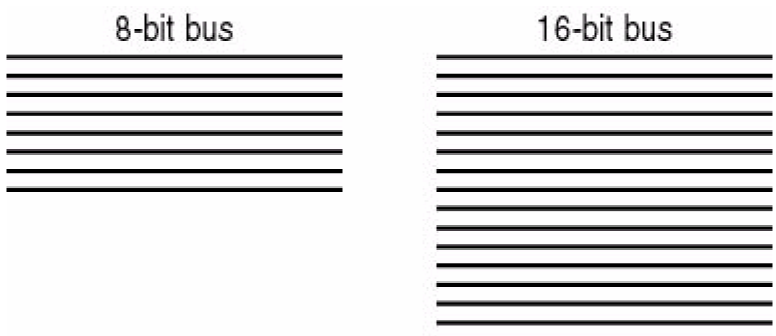
9	00111001	57	E	01100101	101
A	01000001	65	F	01100110	102
B	01000010	66	G	01100111	103
C	01000011	67	H	01101000	104
D	01000010	68	I	0110100	105
E	01000101	69	J	01101010	106
F	01000110	70	K	01101011	107
G	01000111	71	L	01101100	108
H	01001000	72	M	01101101	109
I	01001001	73	N	01101110	110
J	01001010	74	O	01101111	111
K	01001011	75	P	01110000	112
L	01001100	76	Q	01110001	113
M	01001101	77	R	01110010	114
N	01001110	78	S	01110011	115
O	01001111	79	T	01110100	116
P	01010000	80	U	01110101	117
Q	01010001	81	V	01110110	118
R	01010010	82	W	01110111	119
S	01010011	83	X	01111000	120
T	01010100	84	Y	01111001	121
U	01010101	85	Z	01111010	122

Keep in mind that computers are machines, and they do not really perceive numbers as anything other than electrical charges setting a switch on or off. Like binary numbers, electrical charges can exist in only two states—positive or negative. Computers interpret the presence of a charge as one and the absence of a charge as zero. This technology allows a computer to process information.

The Computer Bus

For efficient use of system resources, most communications within a computer need to occur at a much quicker rate than processing signals one at a time would allow. Therefore, the computer moves information through a *bus*. Several types of buses are used within a computer, and they are discussed more fully in later chapters. For now, let's simply look at what a bus is and how it works.

A bus is a group of electrical conductors (usually wires) running parallel to one another that can carry a charge from point a to point b. These conductors can be copper traces on a circuit board or wires in a cable. Usually, they are found in multiples of eight (8, 16, 32, 64, and so on). Early computers used eight conductors for the main system bus, thereby allowing the transmission of eight bits, or one byte, of information at a time. The picture below illustrates an 8-bit and a 16-bit bus.



Computer bus

The physical configuration of a bus isn't as important as its function. A bus provides a common path along which to transmit information in the form of code. It allows any device to receive or send information to any other device on the same bus. This is not unlike the telegraph system in which a single wire was strung from one end of the country to the other. Any town that tapped into the wire could exchange information with any other town also connected to the wire.

Remember: in a computer, a bus is a set of parallel wires or lines to which the CPU, the memory, and all input/output devices are connected. Everything in a computer is connected to a bus. The actual number of wires, or lines, in a bus can vary from one computer to another or even from one part of a computer to another. The bus contains one line for each bit needed to give the address of a device or a location in memory. It also contains one line for each bit of data being transmitted from device to device.

Chapter 3

An Overview of Computers

OBJECTIVES

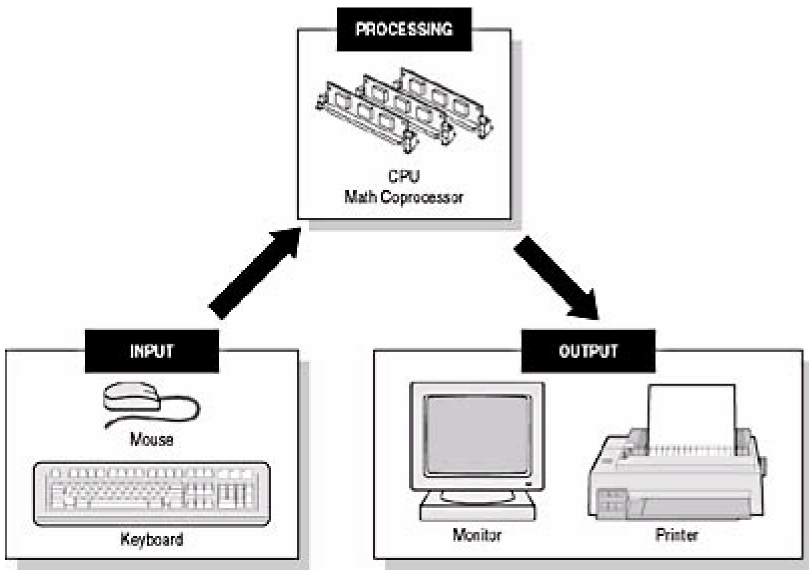
The Three Stages Computing

Components of computers

Support Hardware

The Three Stages of Computing

A modern computer looks like a complicated device. It is constructed of many hardware components connected with what seem to be miles of interwoven wires. Despite this apparent complexity, however, a computer, just like a calculator, handles information in three stages: *input*, *processing*, and *output*. Each piece of hardware can be classified in one (and sometimes two) of these three stages. We can also use these three stages to classify any aspect of a computer's operation or the function of any of its components. During the troubleshooting phase of a repair job, it is often useful to categorize a problem according to which of the three stages it occurs in.



Three stages of computing

Input

Input is the first stage of computing. Input refers to any means that moves data (information) from the outside world into the processor. Today's PC can support a wide variety of input devices; keyboards, mouse devices, voice recognition devices, sound cards, modems, scanners, tape drives, CD/DVD drives, and video cameras are some of the most common.

Processing

Processing is the second stage of computing: the actual manipulation of data by the computer. Computers were designed initially as tools to carry out the tedious task of "number crunching" and then, later, to store large amounts of often redundant data. Today, computers not only fulfill ever-expanding scientific and business roles, but they

also fill our lives with education, entertainment, organization, information processing, and—occasionally—frustration. As we enter this new century, computers have become a necessity of life and are often taken for granted. Even for people who do not own or use a personal computer, they are an increasing part of everyday life. And computer technology hides in many everyday appliances. Microprocessors run most of our mechanical and electronic devices, including cars, cameras, VCRs, microwave ovens, telephones, and the checkout system at the supermarket.

Output

Output is the third stage of computing. All the input and processing in the world won't do us any good unless we can get the information back from the computer in a comprehensible and usable form. Output devices today come in many forms: monitors, printers, fax machines, modems, plotters, CD-recordable discs, sound cards, and more.

Input, Processing, and Output

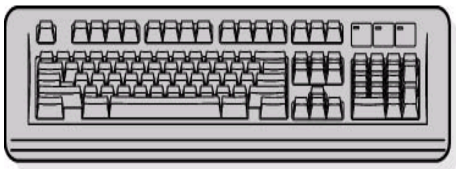
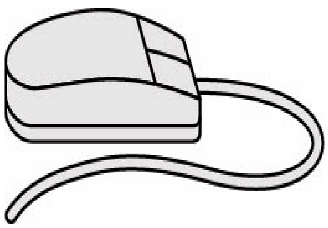
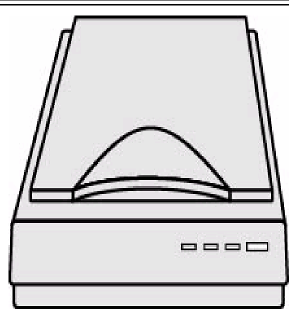
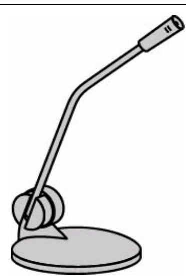
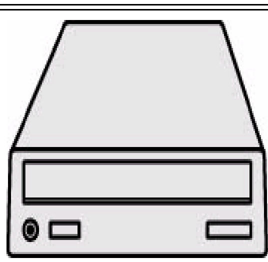
Whenever you sit down at a computer and run an application—whether it is a game, spreadsheet, database, or word processor—you are an active part of the input, processing, and output operation of that computer. The following table provides some examples.

Application	Function
Word processor	Input: Typing your words. Processing: Formatting the text Output: Storing the text and allowing you to retrieve or print it.
Spreadsheet	Input: Typing or providing numbers (such as sales figures). Processing: Applying one or more formulas to the data. Output: Displaying the results of the calculation in numeric or graphical form.
Database	Input: Typing information into a data form. Processing: Indexing and storing the data records. Output: Producing reports showing selected data records.
Game	Input: Moving your chess piece. Processing: Computer calculating how to respond to your move.Output: Computer making a move.

Components of a Computer

Input

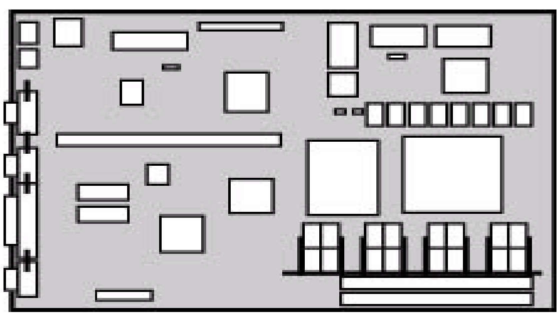
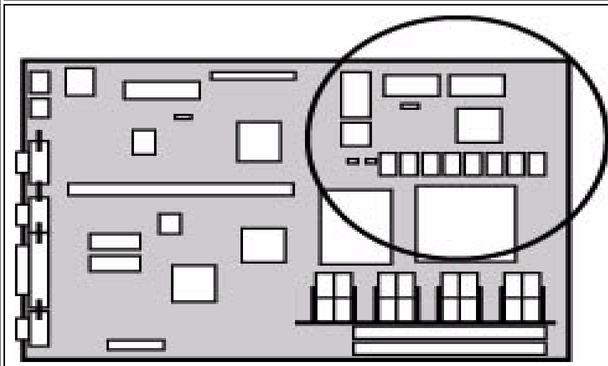
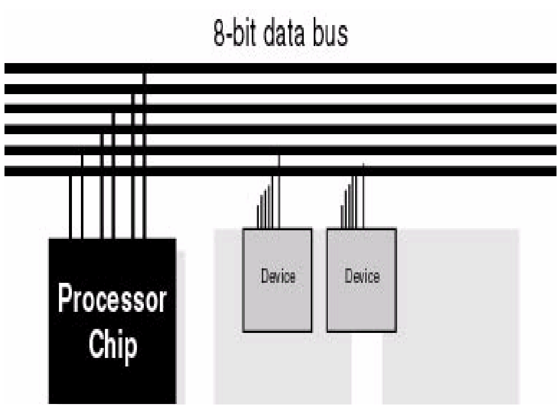
The following table lists the devices that are used to put information into the machine.

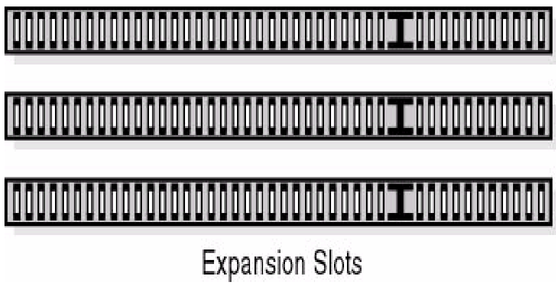
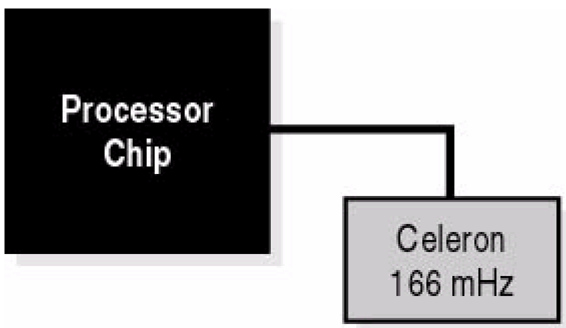
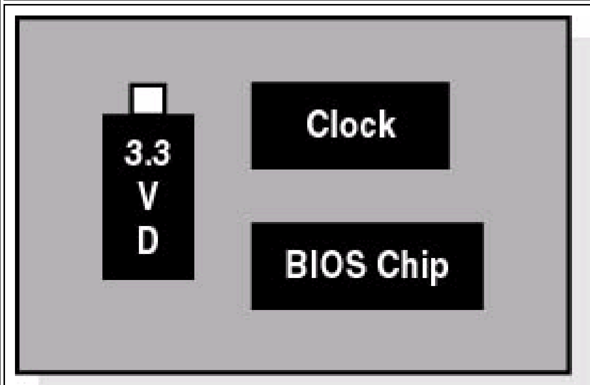
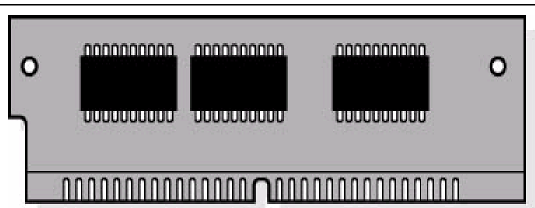
	Device	Description
	**Keyboard**	The primary input device for a computer.
	**Mouse**	Used with graphical interface environments to point and select objects on the system's monitor. Can be purchased in a variety of sizes, shapes, and configurations.
	**Scanner**	Converts printed or photographic information to digital information that can be used by the computer. Works similarly to the scanning process of a photocopy machine.
	**Microphone**	Works like the microphone on a tape recorder. Allows input of voice or music to be converted to digital information and saved to a file.
	**CD-ROM**	Compact disc-read only memory: stores large amounts of data on a compact disc that can be read by a computer.

Processing

The CPU (central processing unit) is the heart and brain of the computer. This one component or "chip" is responsible for all the number crunching and data management. It is truly the centerpiece of any computer. It is so important that whole generations of computer technology are based and measured on each "new and improved" version of the CPU.

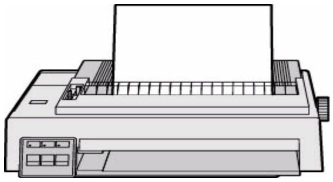
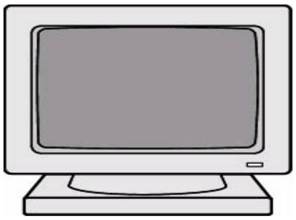
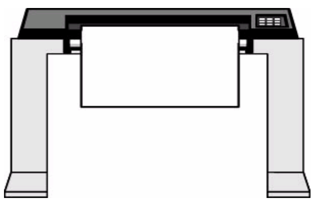
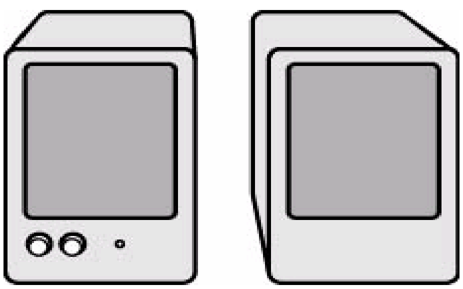
When we refer to the CPU, we are usually speaking of the processor. But the CPU also encompasses several other components that support it with the management of data. These components, when working in harmony, make up the computer we know today. The following table lists these components.

	Component	**Description**
	Motherboard	The large circuit board found inside the computer. Without it, a computer is just a metal box. The motherboard contains all the remaining items in this table; for all practical purposes, it is the computer.
	Chip set	A group of computer chips or integrated circuits (ICs) that, when working together, manage and control the computer system. This set includes the CPU and other chips that control the flow of data throughout the system.
	Data bus	A group of parallel conductors (circuit traces) found on the motherboard and used by the CPU to send and receive data from all the devices in the computer.
	Address bus	A group of parallel conductors (circuit traces) found on the motherboard and used by the CPU to

		"address" memory locations. Determines which information is sent to, or received from, the data bus.
	Expansion slots	Specialized sockets that allow additional devices called expansion cards or, less commonly, circuit boards to be attached to the motherboard. Used to expand or customize a computer, they are extensions of the computer's bus system.
	Clock	Establishes the maximum speed at which the processor can execute commands. Not to be confused with the clock that keeps the date and time.
	Battery	Protects unique information about the setup of the computer against loss when electrical power fails or is turned off. Also maintains the external date and time (not to be confused with the CPU's clock).
	Memory	Stores temporary information (in the form of data bits) that the CPU and software need to keep running.

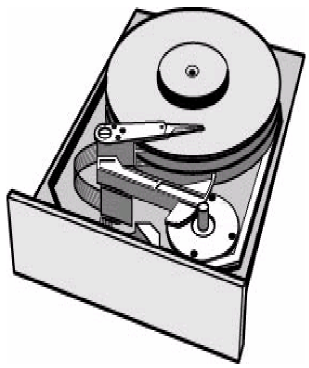
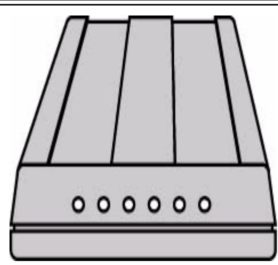
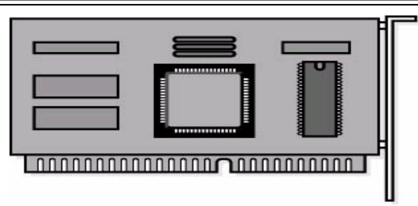
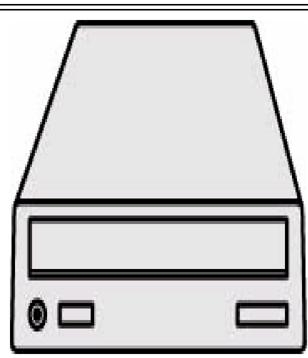
Output

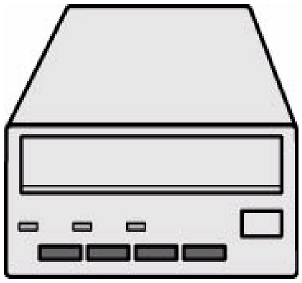
The following table lists some common devices, known as *peripherals*, used exclusively for output.

	Device	Description
	**Printer**	Generates a "hard copy" of information.
	**Monitor**	The primary output device. Visually displays text and graphics.
	**Plotter**	Similar to a printer, but uses pens to draw an image. Most often used with graphics or drawing programs.
	**Speakers**	Reproduce sound. Optional high-quality speakers can be added to provide improved output from games and multimedia software.

Input and Output

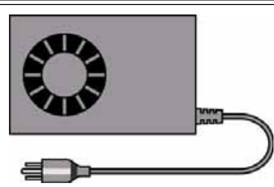
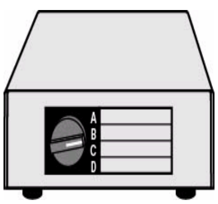
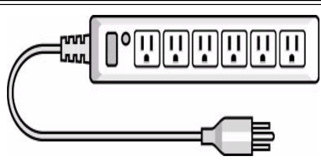
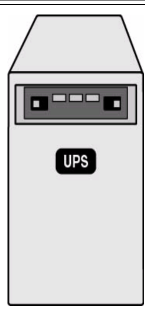
Some devices handle both input and output functions. These devices are called *input/output (I/O)* devices, a term you will encounter quite often.

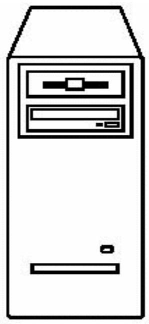
	Device	Description
	**Floppy-disk drive**	Mechanism to read and write to low-capacity, removable, magnetic disks. Used to store and easily transport information.
	Hard-disk drive	High-capacity internal (and sometimes external) magnetic disks for storing data and program files. Also called fixed disks.
	**Modem**	Converts computer data to information that can be transmitted over telephone wires and cable lines. Allows communication between computers over long and short distances.
	**Network card**	An expansion card that allows several computers to connect to each other and share information and programs. Also called network interface card (NIC).
	**CD recorder**	Also called CD/R. You can create a CD with this device, but you can only write to a section of the disc once. Variations on this type of device include CD-RW (CD Read/Write) drives. These products allow you to read, write, and overwrite a special CD-ROM-type disc.

	**Tape drive**	Large-capacity, magnetic, data-storage devices. Ideal for backup and retrieval of large amounts of data. Works like a tape recorder and saves information in a linear format.

Support Hardware

In addition to the devices that support a computer's data-processing functions, there are others that enhance its operation and performance. The following table lists some of these devices.

Figures	Device	Description
	**Power supply**	Converts a local power source (typically 110 volts AC in the U.S.) to 3.3, 5, or 12 volts DC.
	**Switch box**	Allows the user to manually or automatically switch cable connections so that one computer can use several printers or devices with one parallel port.
	**Surge suppressor**	Used to prevent large power spikes (for instance, lightning) from damaging a computer.
	**UPS**	Uninterruptible Power Supply—Acts as both a surge suppresser (to prevent high-power spikes) and a power leveler to provide the computer with a constant source of power. Can even provide power during a power failure or interruption

	**Case**	The box that houses most of the system must provide an environment that minimizes electrical interference to other electronic devices in the area. It should provide a proper heat level for safe operation and bays and connections for drives, circuit boards, and I/O devices.

Chapter 4

The Central Processing Unit

OBJECTIVES

Microprocessors

Replacing & Upgrading Chips

Microprocessors

A *microprocessor* is an integrated circuit that contains a complete CPU on a single chip. In this lesson, we examine the microprocessor from its inception to the current state-of-the-art chip. It is important for a computer technician to understand the development of the processor and what makes each version different from its predecessors. This knowledge gives us an understanding of the enhancements each new design offers over earlier ones and how the system components can take advantage of the new features.

Computer technicians, fortunately, are not required to design microprocessors, only to understand how they work. Microprocessors can be viewed as little black boxes that provide answers or perform a variety of chores on command. We also need to understand the external data bus, because it is the means by which the CPU accesses system resources.

The processor (really a short form for *microprocessor* and also often called the *CPU* or *central processing unit*) is the central component of the PC. The Processor performs the calculations and processing for the PC. The processor plays a significant role in the following important aspects of a computer system:

Performance: The processor determines the PC performance. Although other components also play a key role in determining performance, the processor's capabilities establish the maximum performance of a system.

Software Support: Newer, faster processors enable the use of the latest software

Reliability and Stability: The quality of the processor is one factor that determines how reliably your system will run. While most processors are very dependable, some are not.

Energy Consumption and Cooling: Originally processors consumed relatively little power compared to other system devices. Newer processors can consume a great deal of power. Power consumption has an impact on everything from cooling method selection to overall system reliability.

Motherboard Support: The processor you decide to use in your system will be a major determining factor in what sort of chipset you must use, and hence what motherboard you buy. The motherboard in turn dictates many facets of your system's capabilities and performance.

Processor Manufacturing

When you see a Pentium chip, you are looking at external packaging that is used to house the chip and allow it to interface with the rest of the PC. The actual chip is much smaller and very delicate. It is a square or rectangle of chemically-altered silicon, thinner than a small coin, and with a surface area less than a third of a square inch. There are millions of precisely-placed electronic switches packed into this space.

Cache

It is an intelligent buffer. By using an intelligent algorithm, a cache contains the data accessed most often between a slower peripheral device (RAM) and the faster CPU.

L1 Cache

Level 1 (L1) cache is made directly on the processor die. Therefore, the L1 cache always ran at the full speed of the chip.

L2 Cache

Initially Cache on the motherboard was called the second level, or L2, cache and ran at the slower motherboard speed. Starting with the Pentium Pro and Pentium II, Intel began including L2 cache memory chips directly within the same package as the main processor. Originally, this built-in L2 cache was implemented as physically separate chips contained within the processor package but not a part of the processor die. Because the speed of commercially available cache memory chips could not keep pace with the main processor, most of the L2 cache in these processors ran at one-half speed (Pentium II/III and AMD Athlon), whereas some ran the cache even slower—at two-fifths or even one-third the processor speed (AMD Athlon). The original Pentium II, III, Celeron, and Athlon (Model 1 and 2) processors use 512KB of either one half, two-fifths, or one-third speed L2 cache, as Table 3.1 shows.

L2 Cache Speeds

Processor	**Speed**	**L2 Size**	**L2 Type**	**L2 Speed**
Pentium III	450–600MHz	512KB	External	1/2 core (225–300MHz)
Athlon	550–700MHz	512KB	External	1/2 core (275–350MHz)
Athlon	750–850MHz	512KB	External	2/5 core (300–340MHz)
Athlon	900–1000MHz	512KB	External	1/3 core (300–333MHz)

The Pentium Pro, Pentium II/III Xeon, newer Pentium III, Celeron, K6-3, Athlon (Model 4, codenamed Thunderbird), and Duron processors include full-core speed L2, as shown in Table.

Full-Core Speed Cache

Processor	Speed	L2 Size	L2 Type	L2 Speed
Pentium Pro	150–200MHz	256KB–1MB	External	Full core
K6-3	350–450MHz	256KB	On-die	Full core
Duron	550–700+MHz	64KB	On-die	Full core
Celeron	300–600+MHz	128KB	On-die	Full core
Pentium II Xeon	400–450MHz	512KB–2MB	External	Full core
Athlon	650–1000+MHz	256KB	On-die	Full core
Pentium III	500–1000+MHz	256KB	On-die	Full core
Pentium III Xeon	500–1000+MHz	256KB–2MB	On-die	Full core
Pentium 4	1.3–2.0+GHz	256KB	On-die	Full core
Xeon	1.4–2.0+GHz	256KB–1MB	On-die	Full core

Intel built both the L1 and L2 caches directly on the processor die for the second-generation Celeron, in which, where they both ran at the full-core speed of the chip. The second-generation Pentium III, as well as the AMD K6-3, Athlon, and Duron processors then quickly adopted this type of design. In fact, virtually all-future processors from Intel and AMD have adopted or will adopt on-die L2 cache because it is the only cost-effective way to include the L2 and increase the speed.
Table 3.3 lists the primary specifications for the Intel family of processors used in IBM and compatible PCs.

Table lists the Intel-compatible processors from AMD, Cyrix, NexGen, IDT, and Rise.

Intel Processor Specifications

Processor	CPU Clock	Voltage	Internal Register Size	Data Bus Width	Max. Memory	Level 1 Cache
8088	1x	5v	16-bit	8-bit	1MB	—

8086	1x	5v	16-bit	16-bit	1MB	—
286	1x	5v	16-bit	16-bit	16MB	—
386SX	1x	5v	32-bit	16-bit	16MB	—
386SL	1x	3.3v	32-bit	16-bit	16MB	0KB
386DX	1x	5v	32-bit	32-bit	4GB	—
486SX	1x	5v	32-bit	32-bit	4GB	8KB
486SX2	2x	5v	32-bit	32-bit	4GB	8KB
487SX	1x	5v	32-bit	32-bit	4GB	8KB
486DX	1x	5v	32-bit	32-bit	4GB	8KB
486SL2	1x	3.3v	32-bit	32-bit	4GB	8KB
486DX2	2x	5v	32-bit	32-bit	4GB	8KB
486DX4	2-3x	3.3v	32-bit	32-bit	4GB	16KB
486Pentium OD	2.5x	5v	32-bit	32-bit	4GB	2x16KB
Pentium 60/66	1x	5v	32-bit	64-bit	4GB	2x8KB
Pentium 75-200	1.5–3x	3.3–3.5v	32-bit	64-bit	4GB	2x8KB
Pentium MMX	1.5–4.5x	1.8–2.8v	32-bit	64-bit	4GB	2x16KB
Pentium Pro	2–3x	3.3v	32-bit	64-bit	64GB	2x8KB
Pentium II	3.5–4.5x	1.8–2.8v	32-bit	64-bit	64GB	2x16KB
Pentium II PE	3.5–6x	1.6v	32-bit	64-bit	64GB	2x16KB
Celeron	3.5–4.5x	1.8–2.8v	32-bit	64-bit	64GB	2x16KB
Celeron A	3.5–8x	1.5–2v	32-bit	64-bit	64GB	2x16KB
Celeron III	4.5–9x	1.3–1.6v	32-bit	64-bit	64GB	2x16KB
Pentium III	4–6x	1.8–2v	32-bit	64-bit	64GB	2x16KB
Pentium IIIE	4–9x	1.3–1.7v	32-bit	64-bit	64GB	2x16KB
Pentium II Xeon	4–4.5x	1.8–2.8v	32-bit	64-bit	64GB	2x16KB
Pentium III Xeon	5–6x	1.8–2.8v	32-bit	64-bit	64GB	2x16KB
Pentium IIIE Xeon	4.5–6.5x	1.65v	32-bit	64-bit	64GB	2x16KB

Pentium 4	3–5x	1.7v	32-bit	64-bit	64GB	12+8KB
Pentium 4 Xeon	3–5x	1.7v	32-bit	64-bit	64GB	12+8KB
Itanium	3–5x	1.6v	64-bit	64-bit	16TB	2x16KB

Note

Note in Table 3.3 that the Pentium Pro processor includes 256KB, 512KB, or 1MB of full-core speed L2 cache in a separate die within the chip. The Pentium II/III processors include 512KB of half-core speed L2 cache on the processor card.

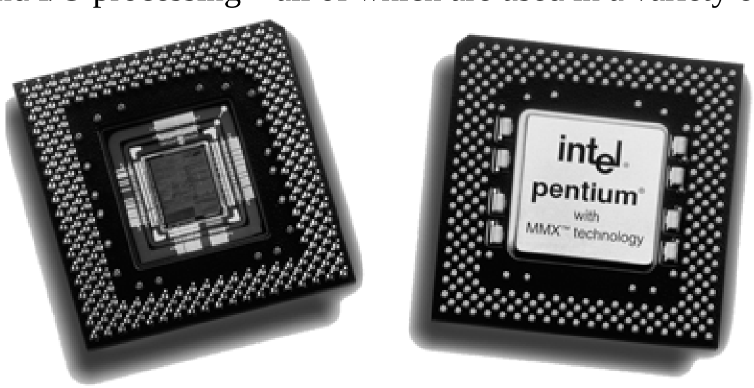
The Celeron, Pentium II PE, and Pentium IIIE processors include full-core speed L2 cache integrated directly within the processor die. The Celeron III uses the same die as the Pentium IIIE, but half of the on-die cache is disabled, leaving 128KB functional.

Data Bus

Perhaps the most common way to describe a processor is by the speed at which it runs and the width of the processor's external data bus. This defines the number of data bits that can be moved into or out of the processor in one cycle. A *bus* is a series of connections that carry common signals. The processor bus discussed most often is the external data bus which is the bundle of wires (or pins) used to send and receive data.

MMX Technology

MMX technology was originally named for multimedia extensions. Intel officially states that it is actually not an abbreviation and stands for nothing other than the letters MMX. MMX technology was introduced in the later fifth-generation Pentium processors as a kind of add-on that improves video compression/decompression, image manipulation, encryption, and I/O processing—all of which are used in a variety of today's software.



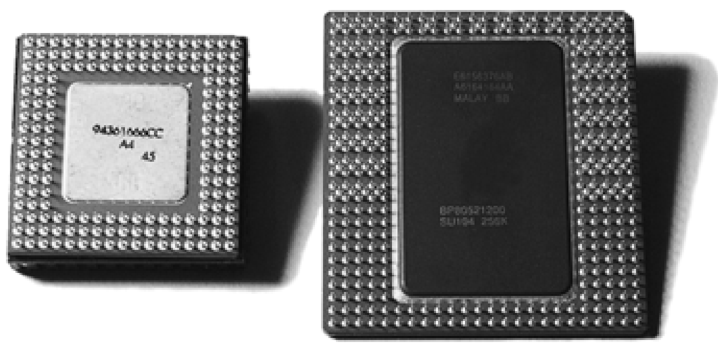
An Intel Pentium MMX chip

Processor Packaging

There are mainly two methods of packaging the processor on the motherboard: -

PGA (Pin-Grid Array) Chip Packaging

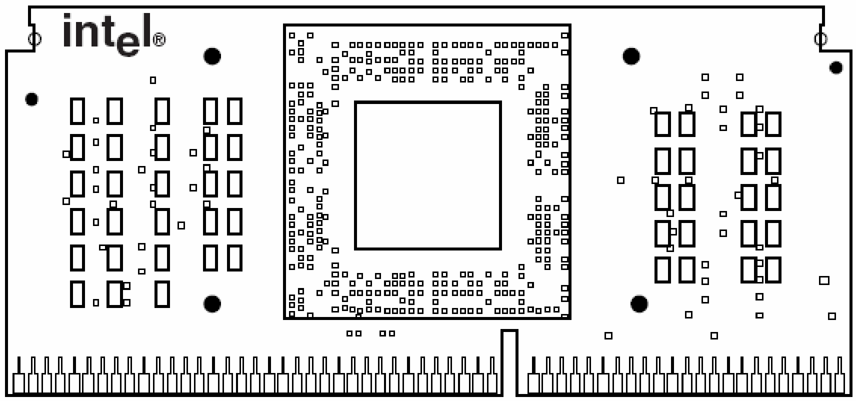
PGA packaging has been the most common chip package used until recently. It was used starting with the 286 processor in the 1980s and is still used today for Pentium processors. PGA chips are inserted into sockets, which are often of a ZIF (zero insertion force) design. A ZIF socket has a lever to allow for easy installation and removal of the chip. Most Pentium processors use a variation on the regular PGA called staggered pin grid array (SPGA), in which the pins are staggered on the underside of the chip rather than in standard rows and columns. This was done to move the pins closer together and decrease the overall size of the chip when a large number of pins are required.



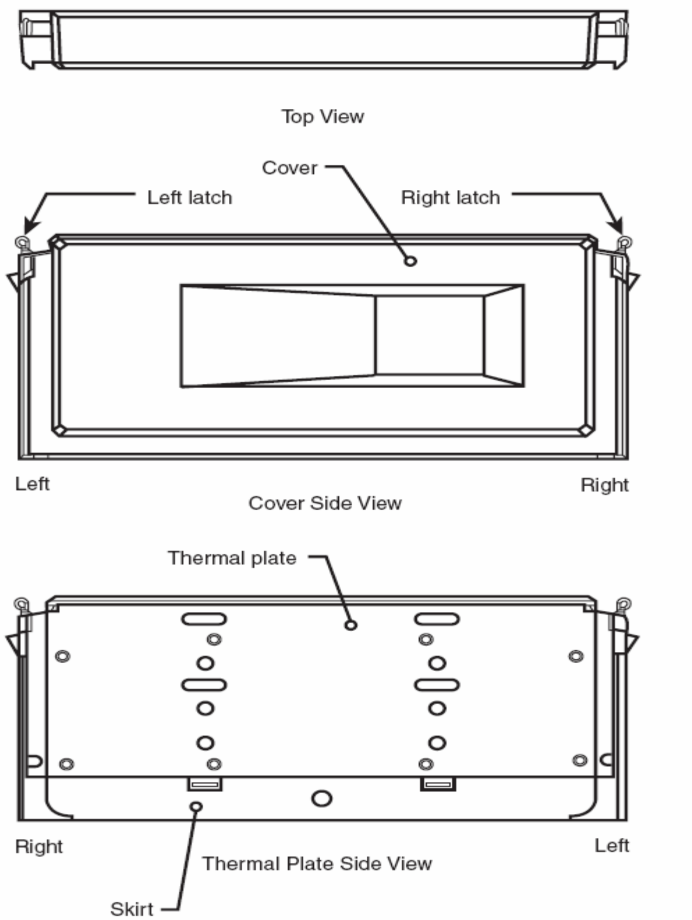
PGA on Pentium 66

Single Edge Contact and Single Edge Processor Packaging

Pentium II/III chips are characterized by their single edge contact (SEC) cartridge design. The processor, along with several L2 cache chips, is mounted on a small circuit board (much like an oversized memory SIMM), which is then sealed in a metal and plastic cartridge. The cartridge is then plugged into the motherboard through an edge connector called Slot 1, which looks very similar to an adapter card slot.



Pentium II Processor Slot 1



Pentium II Processor SEC package parts.

Processor Sockets and Slots

Intel and AMD have created a set of socket and slot designs for their processors. Each socket or slot is designed to support a different range of original and upgrade processors. The table below shows the specifications of these sockets.

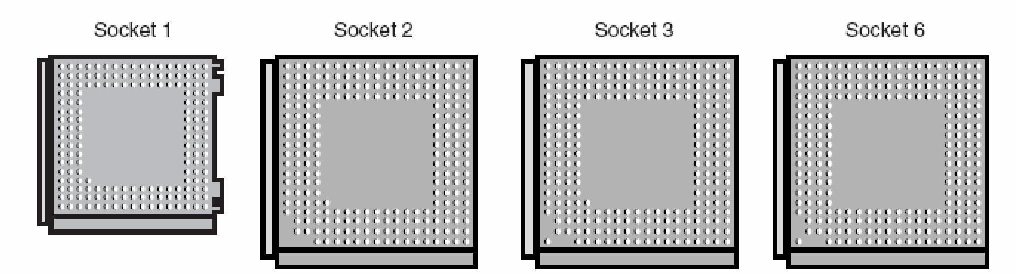
CPU Socket and Slot Types and Specifications

Socket Number	Pin	Pin Layout	Voltage	Supported Processors
Socket 1	169	17x17 PGA	5v	486 SX/SX2, DX/DX2, DX4 OverDrive,
Socket 2	238	19x19 PGA	5v	486 SX/SX2, DX/DX2, DX4 OverDrive, 486 Pentium Overdrive
Socket 3	237	19x19 PGA	5v/3.3v	486 SX/SX2, DX/DX2, DX4 OverDrive, 486 Pentium OverDrive, AMD5x86
Socket 4	273	21x21 SPGA	5v	Pentium 60/66, OverDrive
Socket 5	320	37x37 SPGA	3.3/3.5v	486 SX/SX2, DX/DX2, DX4 OverDrive,
Socket 6*	**235**	**19x19 PGA**	**3.3v**	**486 DX4, 486 Pentium Overdrive**
Socket 7	321	37x37 SPGA	VRM	Pentium 75–233+, MMX, OverDrive, AMD K5/K6, Cyrix M1/II
Socket 8	387	DP-SPGA	Auto VRM	Pentium Pro, Overdrive
Socket 370	370	37x37 SPGA	Auto VRM	Celeron/Pentium III PPGA/FC-PGA
Socket PAC418	418	38x22 split SPGA	Auto VRM	Itanium
Socket 423	423	39x39 SPGA	Auto VRM	Pentium 4
Socket A (462)	462	37x37 PGA	Auto VRM	AMD Athlon/Duron FC-PGA
Socket 603	603	31x25 SPGA	Auto VRM	Xeon (P4)
Slot A	242	Slot	Auto VRM	AMD Athlon SECC
Slot 1	242	Slot	Auto	Pentium II/III, Celeron

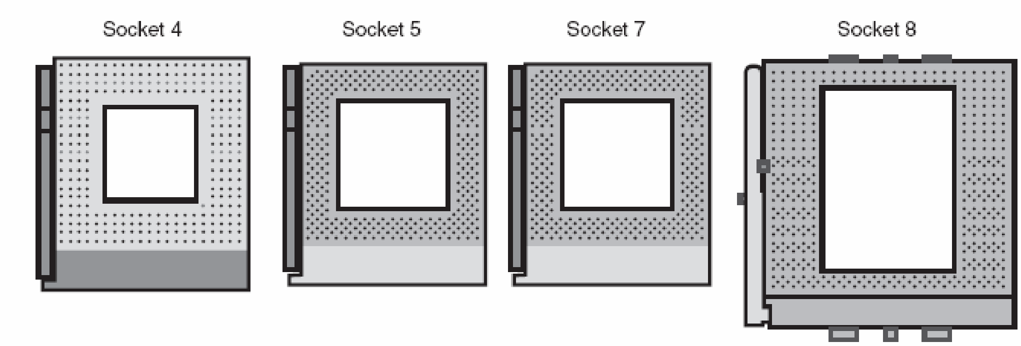
(SC242)			VRM	SECC
Slot 2 (SC330)	330	Slot	Auto VRM	Pentium II/III Xeon SECC

*Socket 6 was a paper standard only and was never actually implemented in any systems.
PGA = Pin grid array
PPGA = Plastic pin grid array
FC-PGA = Flip chip pin grid array
SPGA = Staggered pin grid array
SECC = Single edge contact cartridge
PAC = Pin array cartridge
VRM = Voltage regulator module

Sockets 1, 2, 3, and 6 are 486 processor sockets and are shown together in Figure 3.10 so you can see the overall size comparisons and pin arrangements between these sockets. Sockets 4, 5, 7, and 8 are Pentium and Pentium Pro processor sockets and are shown together in Figure 3.11 so you can see the overall size comparisons and pin arrangements between these sockets. More detailed drawings of each socket are included throughout the remainder of this section with the thorough descriptions of the sockets.



486 processor sockets.



Pentium and Pentium Pro processor sockets.

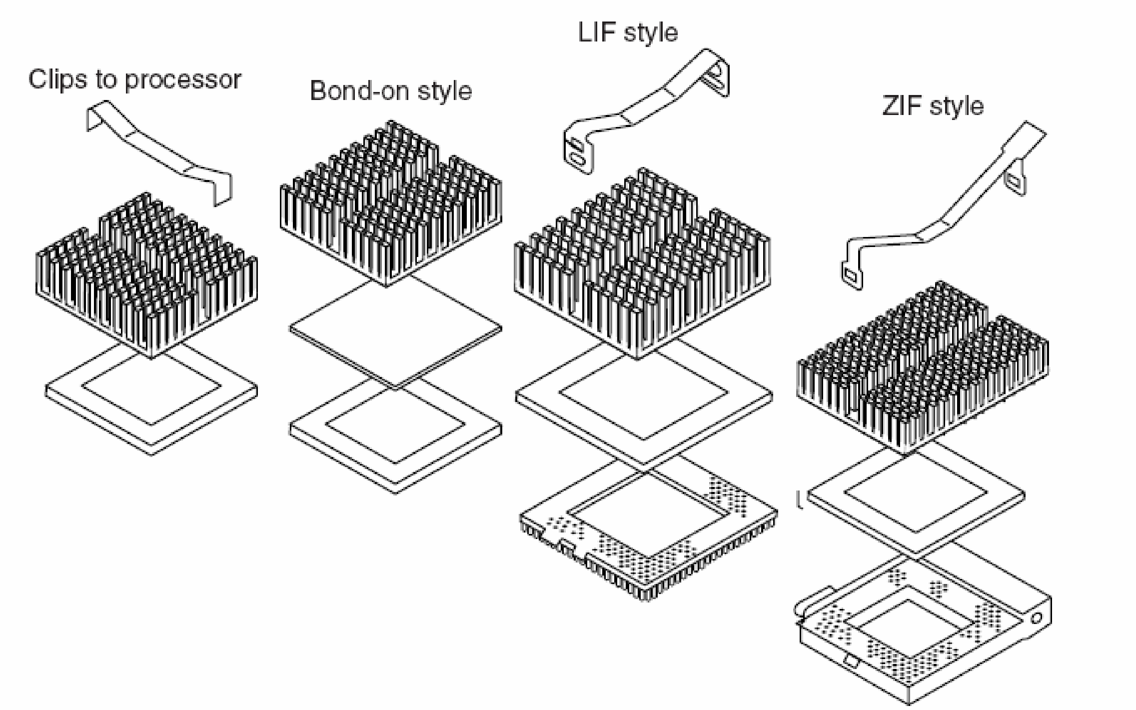
Heat and Cooling Problems

The higher-speed processors typically consume more power and therefore generate more heat. The processor is usually the single most power hungry chip in a system, and in most

situations, the fan inside your computer case might not be capable of handling the load without some help.

Heat sinks

To cool a system in which processor heat is a problem, you can buy (for less than $5, in most cases) a special attachment for the CPU chip called a *heat sink*, which draws heat away from the CPU chip. It does this by using a thermal conductor (usually metal) to carry heat away from the processor into fins that expose a high amount of surface area to moving air. This enables the air to be heated, thus cooling the heat sink and the processor as well. Just like the radiator in your car, the heat sink depends on airflow. With no moving air, a heat sink is incapable of radiating the heat away. The heat sink is clipped or glued to the processor. A variety of heat sinks and attachment methods exist. The picture below shows various passive heat sinks and attachment methods.

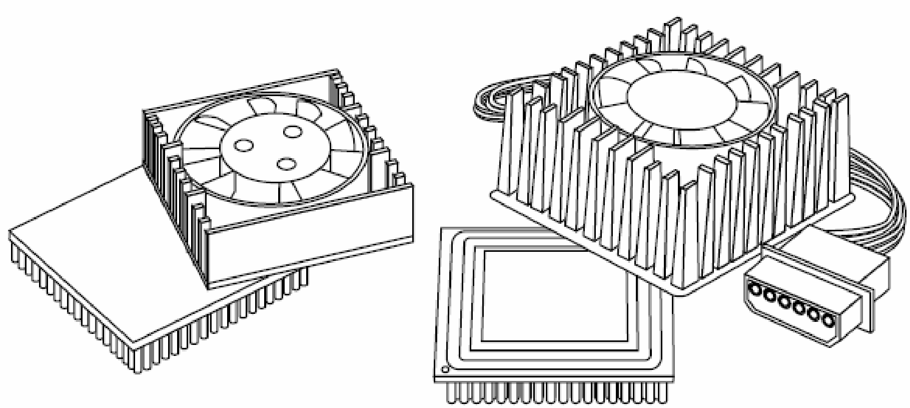


Passive heat sinks for socketed processors showing various attachment methods.

Warning

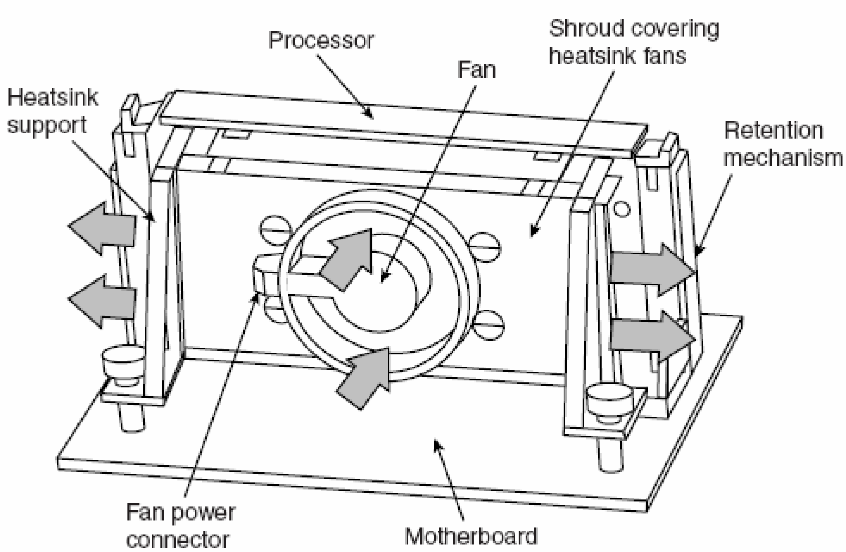
According to data from Intel, heat sink clips are the number-two destroyer of motherboards (screwdrivers are number one).
When installing or removing a heat sink that is clipped on, be sure you don't scrape the surface of the motherboard. In most cases, the clips hook over protrusions in the socket, and when installing or removing the clips, scratching or scraping the surface of the board right below where the clip ends attach is very easy. I like to place a thin sheet of plastic underneath the edge of the clip while I work, especially if board traces that can be scratched are in the vicinity.

The active heat sinks for the processors have a fan built into them.



Active heat sinks for socketed processors.

This picture shows an active heat sink arrangement on a Pentium II/III type processor. This is common on what Intel calls its "boxed processors.



Processor Upgrades

There is a possibility of upgrading the processor to the next higher speed. For example, if you have a motherboard with Socket 3, you can put virtually any 486 processor in it; if you have a Socket 7 motherboard, it should be capable of accepting virtually any Pentium processor. Before you upgrade the processor on a particular motherboard, you have to examine the motherboard documentation to check for the compatibility of following main concerns: -

- Processor Socket
- Voltage
- BIOS

For example, if your motherboard has a Pentium Socket 5, you can install a Pentium MMX 233MHz processor with a 2.8v voltage regulator adapter, or optionally an AMD-K6, also with a voltage regulator adapter. If you have Socket 7, your motherboard should be capable of supporting the lower voltage Pentium MMX or AMD-K6 series directly without any adapters. The K6-2 and K6-3 are the fastest and best processors for Socket 7 motherboards.

OverDrive Processors

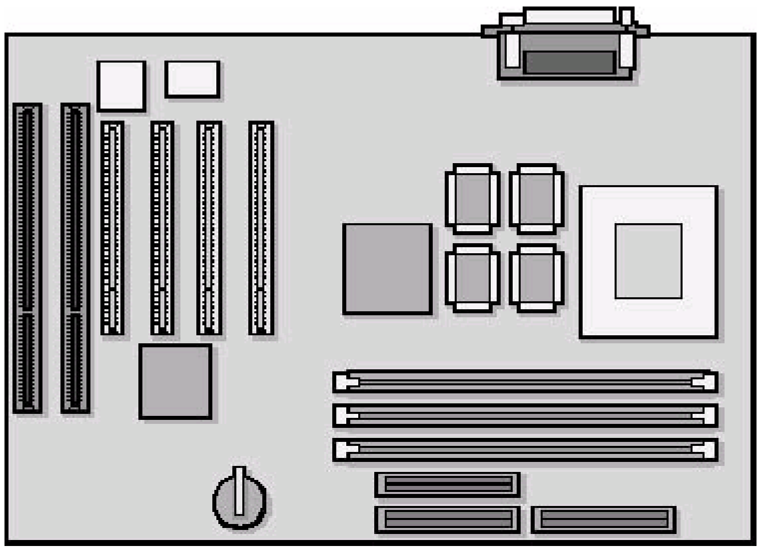
Intel at one time offered special OverDrive processors for upgrading systems. Normally, these were repackaged versions of the standard processors, sometimes including necessary voltage regulators and fans. They have all been withdrawn, and Intel has not announced any new versions.

The External Data Bus

Information is transmitted throughout a computer by binary code traveling through a bus. The *external data bus* (also known as the external bus or simply data bus) is the primary route for data in a PC. All data-handling components or optional data devices are connected to it; therefore, any information (code) placed on that bus is available to all devices connected to the computer.

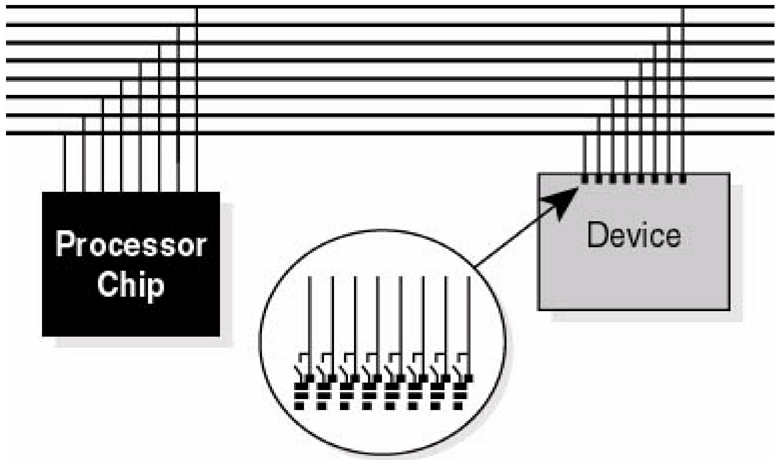
Early computers used eight conductors (an 8-bit data bus), which allowed for the transfer of 1 byte of information at a time. As computers evolved, the width of the external data bus increased to 16, 32, and finally to the current width of 64 conductors. The wider bus lets more data flow at the same time, just as adding more lanes to a highway allows more cars to move through a point in a given amount of time.

The picture below shows a CPU attached to its motherboard. The motherboard is the main circuit board, which contains the external data bus and connection for expansion devices that are not part of the board's basic design. The expansion slots act as "on ramps" to the external bus. Expansion cards, once commonly known as "daughter cards," are placed in slots on the motherboard. Other forms of onramp are the slots that hold memory or the sets of pins used to attach drive cables. Connectors on the motherboard grant access to the data bus for keyboards, mouse devices, and peripheral devices like modems and printers through the use of COM and LPT ports.



Motherboard

To understand how a computer moves data between components, visualize each device on the data bus (including the CPU) connected to the bus by means of a collection of on/off switches. By "looking at" which conductors have power and which ones do not, the device can read the data as it is sent by another device. The on-off state of a line gives the value of 0 (on) or 1 (off). The wires "spell out" a code of binary numbers that the computer interprets and then routes to another system component or to the user by means of an output device such as a monitor or printer. Communication occurs when voltage is properly applied to, or read from, any of the conductors by the system. The picture below illustrates a data bus connected to a CPU and a device.

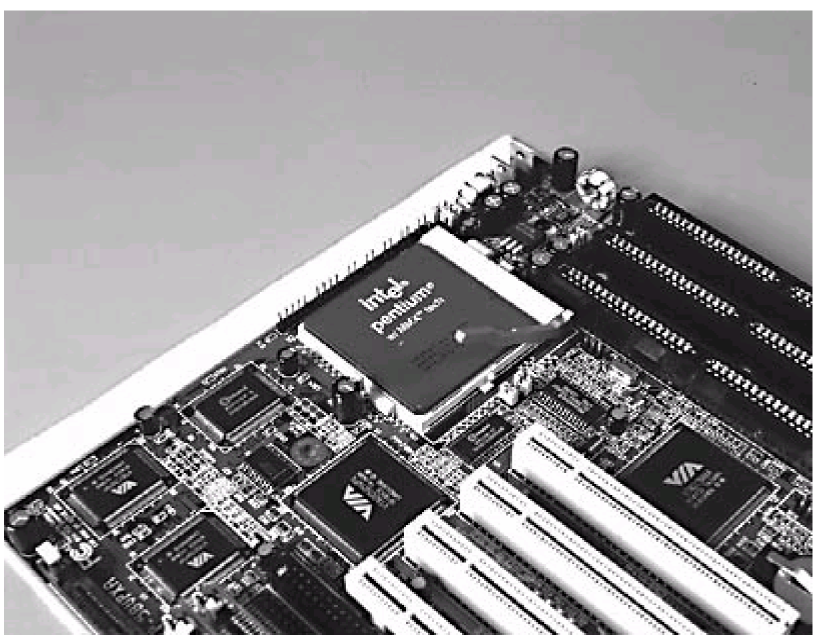


External data bus

Coded messages can be sent into or out of any device connected to the external data bus. Think of the data bus as a large highway with parallel lanes. Extending that analogy, bits are like cars traveling side by side each carries part of a coded message. Microprocessors are used to turn the coded messages into data that performs a meaningful task for the computer's user.

The CPU

The CPU is the part of a computer in which arithmetic and logical operations are performed and instructions are decoded and executed. The CPU controls the operation of the computer. Early PCs used several chips to handle the task. Some functions are still handled by support chips, which are often referred to collectively as a "chip set." The picture below shows a close-up of a CPU and other chips on a motherboard.



CPU

Transistors

Transistors, the main components of microprocessors, are small, electronic switches. The on/off positions of the transistors form the binary codes discussed earlier in this lesson. Although transistors might seem simple, their development required many years of painstaking research. Before transistors were available, computers relied on slow, inefficient vacuum tubes and mechanical switches to process information. The first large-scale computers took up a huge amount of space, and technicians actually went inside them to "program" by turning on and off specific tubes!

Many materials, including most metals, allow electrical current to flow through them these are known as electrical conductors. Materials that don't pass electrical current are called insulators. Pure silicon (which is used to make most transistors) is a *semiconductor*; its degree of conductivity can be adjusted, or modulated, by adding impurities during production.

Transistor switches have three terminals: the source, the gate, and the drain. When positive voltage is applied to the gate, electrons are attracted, forming an electron channel between the source and the drain. Positive voltage applied to the drain pulls electrons from the source to the drain, turning the transistor on. Removing the voltage turns it off by breaking the pathway.

Integrated Circuits

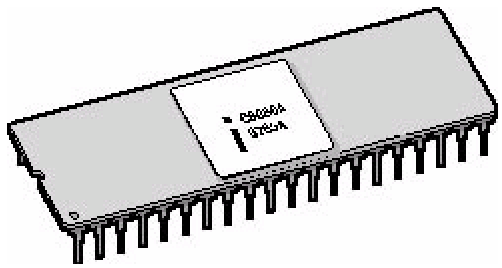
An *integrated circuit (IC)* is an electronic device consisting of a number of miniature transistors and other circuit elements (resistors and capacitors, for instance). An IC functions just as a large collection of these parts would, but it is a fraction of the size and uses a fraction of the power. ICs make today's microelectronics possible. The original transistors were small plastic boxes about the size of a peanut (outside its shell), and could handle only one function. The word "integrated" denotes that IC devices combine many circuits and some of their functions into one package. A prime example of this technology is the microprocessor.

Microprocessors

On November 15, 1971, Intel shipped the commercial microprocessor, Model 4004. It ran a product called the Busicom calculator. The 108-KHz 4004 had 2300 transistors and a 4-bit data bus and could address 640 bytes of RAM. Computer engineers quickly took advantage of the potential this new type of chip offered, leading the way to the first personal computers.

A year later, the Intel 8008 appeared. *Radio Electronics Magazine* reported that hobbyist Don Lancaster used an 8008 to build what is considered the first personal computer. The article called it a "TV typewriter."

The Intel 8080 appeared in 1974. It sold then for $400, and now sells for about one dollar. It powered traffic lights, but of more interest to our discussion is the fact that it formed the core of the Altair computer of 1975. It was sold in kit form for $395 and was named for a world in the Star Trek TV series. The picture below shows a picture of the 8080 die. By today's standards, it was very weak: 6000 transistors, an 8-bit bus, and a 2-MHz clock speed. It could address 64 K of RAM, and users programmed the Altair by throwing manual switches located on the case.



The Intel 8080 Microprocessor

Microprocessor Design

Before going further into microprocessor-development history, it is important to discuss in general terms how they operate. Microprocessors are usually divided into three subsystems: the control unit (CU), the arithmetic logic unit (ALU), and the input/output unit. The term CPU is used to denote a combined CU and ALU, contained in a single package.

The advent of the control unit marked a radical improvement in processor design, allowing CPU operations to be based in part on code provided by an external program like a BIOS (basic input/output system). This extended the ability of a PC to use new hardware components that were not part of the original design.

The ALU is just what its name implies—the part of the IC that handles the basic, math functions of computation. The I/O unit fetches data from the outside and passes data back to the external bus.

Registers

Registers are temporary memory storage areas used during data manipulation. Physically, registers are rows of microscopic switches, which are set on or off. Each row forms a binary number: off = 0, on = 1. Hence (reading from right to left) off.off.on equals the

number 1. Off.on.on equals the number three (0+2+1). The CPU uses registers like scratch pads, to hold data while it works on a task. Changes in data during an operation are also stored in a register, then sent out to other components as the job is finished. The number and width of a register varies from one type of machine to another. The wider the register, the more bits the machine can handle at one time—just as with the width of the external bus. As register width moved from 4 to 8 to 16 to 32 to 64 bits, PCs increased in performance.

Codes

Computers use various binary-based codes to represent information. We saw how ASCII code is a binary representation of characters on a keyboard. These codes are sent on the external data bus by a system component to be read by other devices. Press a key on a PC keyboard and an ASCII code is generated and sent over the data bus. Transferring information to and from the CPU (and other hardware) is only the first step in manipulating data.

Other codes tell the PC how to display data on the monitor, talk to devices such as printers, and take in data streams from scanners. Each of those operations requires system resources and the manipulation of binary numbers.

In addition to the code that requires data, special *machine code* is required in order for the CPU to turn the string of numbers into something useful to an application. As with the data code, this machine code is sent in the form of binary numbers on the data bus. The CPUs in turn are different enough that a code system must be written specifically for each of them.

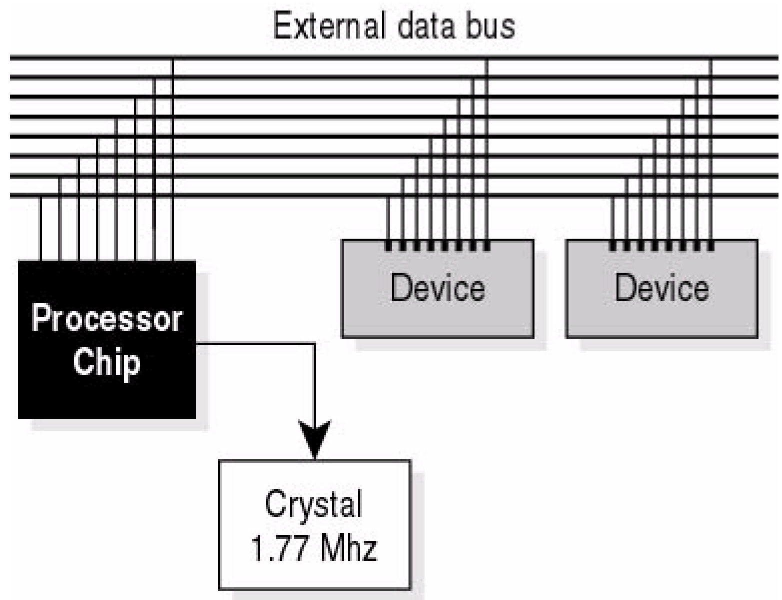
The Clock

Timing is essential in PC operations. Without some means of synchronization, chaos would ensue. Timing allows the electronic devices in the computer to coordinate and execute all internal commands in the proper order.

Timing is achieved by placing a special conductor in the CPU and pulsing it with voltage. Each pulse of voltage received by this conductor is called a "clock cycle." All the switching activity in the computer occurs while the clock is sending a pulse. This process somewhat resembles several musicians using a metronome to synchronize their playing, with all the violinists moving their bows at the same time. Thanks to this synchronization, you get musical phrasing instead of a jumble of notes.

Virtually every computer command needs at least two clock cycles. Some commands might require hundreds of clock cycles to process. The picture below shows an external

data bus with a CPU and two devices. Notice that the crystal or clock is attached to the CPU to generate the timing.



CPU with clock

Clock Speed

It is common for computers to be marketed to consumers based on features that show off their best points. One principal selling point is the system clock rate—measured in megahertz (MHz) or millions of cycles per second. The clock rate suggests how many commands can be completed in two cycles (the minimum time required to execute a command). The process of adding two numbers together would take about four commands (eight clock cycles). A computer running at 450 MHz can do about 44 million simple calculations per second.

Clock speed is determined by the CPU manufacturer and represents the fastest speed at which the CPU can be reliably operated. The Intel 8088 processor, as used in the original IBM PC, had a clock speed of 4.77 MHz. Today's processors have clock speeds that run up to and, in some, exceed 2.5 GHz.

The system crystal determines the speed at which a CPU operates. The system crystal is usually a quartz oscillator, very similar to the one in a wristwatch. You can find the system crystal soldered to the motherboard. Look for a silver part, usually with a label that indicates the crystal speed.

IMPORTANT

A computer has two clocks: one to set the speed and timing and a second clock to keep time for date/time calculations. They are two entirely different devices.

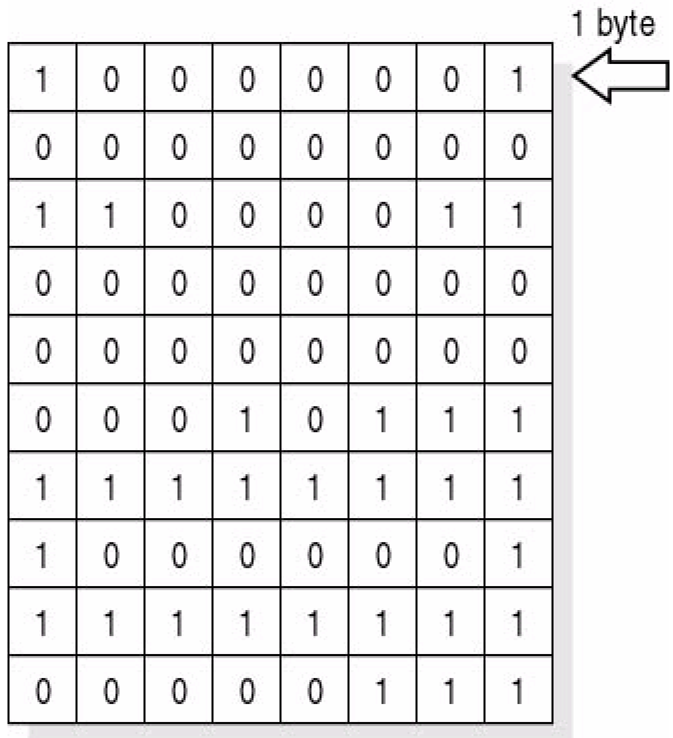
Memory

The CPU's ability to hold large amounts of information at once is very limited. To compensate, additional chips are installed in the computer for the sole purpose of temporarily storing information that the CPU needs. These chips are called random access memory (RAM). The term *random access* is used because the CPU can place or retrieve bytes of information in or from any RAM *location* at any time.

Address Bus

The word "location" is italicized in the last paragraph to underscore the importance of location in PC memory operations. The content of RAM is changing all the time, as programs and the computer itself use portions of it to note, calculate, and hold results of actions. It is essential for the system to know what memory is assigned to which task and when that memory is free for a new use. To do so, the system has to have a way to address segments of memory and to quickly change the holdings in that position. The portion of the PC that does this is the *address bus*.

Think of the address bus as a large, virtual table in which the columns are individual bits (like letters) and each row contains a string of bits (making up a word). The actual lengths of these words will vary depending on the number of bits the address bus can handle in a single pass. The picture below shows a table containing 1s and 0s. Each segment is given an address, just like the one that identifies a home or post office box. The system uses this address to send data to or retrieve data from memory.



Memory spreadsheet

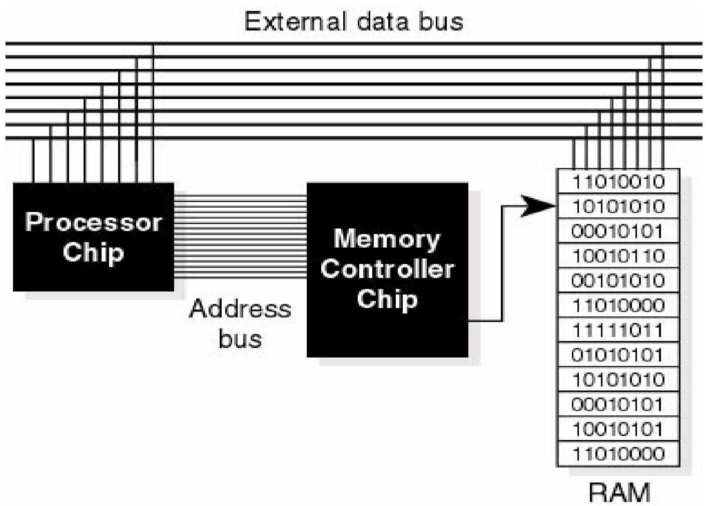
Like all the other buses in a PC, this one is a collection of conductors. It links the physical memory to the system and moves signals as memory is used. The number of conductors in the address bus determines the maximum amount of memory that can be used (memory that is *addressable*) by the CPU. Remember that computers count in binary notation. Each binary digit—in this case, a conductor—that is added to the left will double the number of possible combinations.

Early data buses used eight conductors and, therefore, 256 (2^8) combinations of code where possible. The maximum number of patterns a system can generate determines how much RAM the data bus can address. The 8088 used 20 address conductors and could address up to 1,048,576 bytes of memory locations, or 2^{20}. Today's PCs can address a lot more than that, and, in many cases, the actual limiting factor is not the number of patterns, but the capacity of the motherboard to socket memory chips. In all cases, the total amount of memory is the factor of 2^X, where X = the number of connectors.

The CPU does not directly connect to the memory bus, but sends requests and obtains results using the system's memory controller. This circuitry acts as both postmaster and translator, providing the proper strings of data in the right order, at the right time, and in a form the CPU can use. As mentioned before, any write or read action will require at least two clock cycles to execute. (It can require more clock cycles on systems that do not have memory tuned to the maximum system clock speed. In that case, the PC will have to use

additional clock cycles while it waits for the memory to be ready for the next part of the operation.)

The picture below shows a diagram of the process with the CPU and RAM stack on the external data bus. The address bus is connected to the memory controller. It fetches and places data in memory.



CPU and RAM

How Microprocessors Work

Current CPUs, such as the Intel Pentium III, & PIV are collections of millions of switches and bus pathways. They operate all kinds of machines, in addition to PCs, and are found in cameras, cars, microwave ovens, TVs, and all kinds of gadgets. Here, however, we are interested only in how they work inside a PC. Let's look at a simple task: adding two numbers such as 2 and 2 together and obtaining their sum (2 + 2 = 4). The CPU can do math problems very quickly, but it requires several very quick steps to do it. Knowing how a CPU performs a simple task will help you understand how developments in PC design have improved PC performance.

When the user pushes a number key (in a program like Calculator, which can add numbers), the keystroke causes the microprocessor's prefetch unit to ask for instructions on what to do with the new data. The data is sent through the address bus to the PC's RAM and is placed in the instruction cache, with a reference code (let's call it 2 = a).

The prefetch unit obtains a copy of the code and sends it to the decode unit. There it is translated into a string of binary code and routed to the control unit and the data cache to

tell them what to do with the instruction. The control unit sends it to an address called "X" in the data cache to await the next part of the process.

When the plus (+) key is pressed, the prefetch unit again asks the instruction cache for instructions about what to do with the new data. The prefetch unit translates the code and passes it to the control unit and data cache, which alerts the ALU that an ADD function will be carried out. The process is repeated when the user presses the "2" key.

Next (yes, there's still more to do), the control unit takes the code and sends the actual ADD command to the ALU. The ALU sums "a" and "b" are added together after they have been sent up from the data cache. The ALU sends the code for "4" to be stored in an address register.

Pressing the equal sign (=) key is the last act the user must execute before getting the answer, but the computer still has a good bit of work ahead of it. The prefetch unit checks the instruction cache for help in dealing with the new keystroke. The resulting instruction is stored, and a copy of the code is sent to the decode unit for processing. There, the instruction is translated into binary code and routed to the control unit. Now that the sum has (finally) been computed, a print command retrieves the proper address, registers the contents, and displays them. (That involves a separate flurry of activity in the display system, which we won't worry about.)

As you can see, a microprocessor must go through many more steps than human beings are required to take, merely to arrive at the conclusion that 2 + 2 = 4. The computer must execute a complicated dance in order to manage the code, place it, and fetch it in memory; then it has to be told what to do with it. Yet the result usually appears as fast as you can type the request. You can see that clock cycles and, hence, processor speed, have a significant effect on performance. Other issues that affect performance include memory access and speed, as well as the response time of components such as the display system.

PC Microprocessor Developments and Features

PC microprocessor design grows more complex with each generation, and CPU packaging keeps changing to provide room for additional features and operating requirements. Microprocessors have evolved from the 4004 described earlier into today's high-speed Pentiums. Each new processor has brought higher performance and spawned new technology. Six basic elements are customarily used to gauge the performance and capability of a CPU design:

- Speed: The maximum number of clock cycles measured in megahertz. The higher the speed, the quicker a command will be executed.
- Number of transistors: More switches, more computing power.

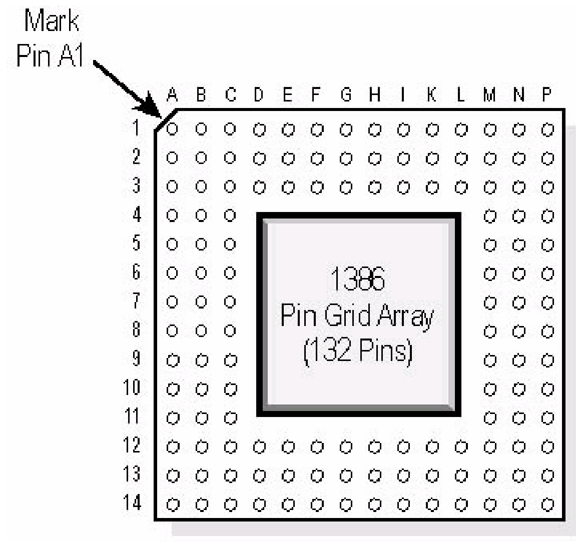
- Registers: The size (in bits) of the internal registers. The larger the registers, the more complicated the commands that can be processed in one step.
- External data bus: As data bus size increases, so does the amount and complexity of code (information) that can be transferred between all devices in the computer.
- Address bus: The size of the address bus determines the maximum amount of memory that can be addressed by the CPU.
- Internal cache: The internal cache is high-speed memory built into the processor. This is a place to store frequently used data instead of sending it to slower devices (speed is relative in computers) such as RAM and hard disk drives. It is built into the processor and has a dramatic effect on speed. We cover cache in more detail later in this lesson.

Intel has held most of the PC CPU market share since the original IBM PC was introduced. Closely following each new Intel launch, rivals such as Advanced Microdevices (AMD) and Cyrix have offered alternative chips that are generally compatible with the Intel models. This development, in turn, drives prices down and spurs a new round of CPU design. Another player is Motorola, a firm that manufactures the microprocessors used in the Apple family of computers, among others.

Virtual Memory

Virtual memory is the art of using hard disk space to hold data not immediately required by the processor; it is placed in and out of RAM as needed. Although using virtual memory slowed the system down (electronic RAM is much faster than a mechanical hard drive), it allowed the 286 to address up to 1 GB (gigabyte—one thousand megabytes) of memory (16 MB of actual memory and 984 MB of virtual memory). Virtual memory required the use of operating systems more advanced than MS-DOS, leading to the development of products such as Microsoft Windows, IBM OS/2, and SCO's PC version of UNIX.

The original 286 processor came packaged in DIP (already shown), PGA (pin grid array), and PLCC (plastic leadless chip carrier) designs. The PLCC can be recognized by the arrangement of thin legs around its perimeter. The PLCC's major advantage is its stronger leads (pins), which make it more difficult to damage during removal or installation. PLCCs became popular because they made it easier to upgrade a PC with a faster CPU.



Plastic Leadless Chip Carrier (PLCC) CPU Package

System Memory Management

System Memory Management (SMM) is a hardware-based function that allows the microprocessor to selectively shut down the monitor, hard drives, and any other peripherals not in use. SMM works at the chip level; the microprocessor can be operating in real, protected, or virtual 8086 mode. SMM is transparent to all software running on the system, which decreases the likelihood of lockups.

Vendors such as AMD rode the wave with their own editions of the 486 for users feeling a need for greater speed. The following table lists the most popular 486 chips and third-party work-alikes.

Chip	CPU Speed (MHz)	Register Width	External Data Bus	Address Bus	Internal Cache
Intel 80486DX	25, 33, 50	32-bit	32-bit	32-bit	8 KB
Intel 80486DX/2	50, 66	32-bit	32-bit	32-bit	8 KB
Intel 80486DX/4	75, 100	32-bit	32-bit	32-bit	16 KB
Intel 80486SX	16, 20, 25	32-bit	32-bit	32-bit	8 KB
Intel 80486SL	16, 20, 25	32-bit	32-bit	32-bit	8 KB

AMD AM486DX	33, 40	32-bit	32-bit	32-bit	8 KB
AMD AM486DXLV	33	32-bit	32-bit	32-bit	8 KB
AMD AM486DX2	50, 80	32-bit	32-bit	32-bit	8 KB
AMD AM486DX4	100, 120	32-bit	32-bit	32-bit	8 KB
AMD AM486DX "Enhanced"	120, 133	32-bit	32-bit	32-bit	16 KB W/B
AMD AM486DXL2	50, 80	32-bit	32-bit	32-bit	8 KB
AMD AM486SX	33, 40	32-bit	32-bit	32-bit	8 KB
AMD AM486SXLV	33	32-bit	32-bit	32-bit	8 KB
AMD AM486SX2	33	32-bit	32-bit	32-bit	8 KB
CYRX CX486DX	33	32-bit	32-bit	32-bit	8 KB W/B
CYRX CX486DX2	50-80	32-bit	32-bit	32-bit	8 KB W/B
CYRX CX486DLC	33-40	32-bit	32-bit	32-bit	1 KB W/B
CYRX CX486SLC	20-33	32-bit	32-bit	32-bit	1 KB W/B
CYRX CX486SLC2	50	32-bit	32-bit	32-bit	1 KB W/B

Heat Sinks and Fans

The 486 is notable for one other item, the addition of a standard heat sink and, usually, a fan mounted on the CPU and powered by the PC. To maintain stable operation, the PC must provide proper cooling for the 486 and newer CPUs. Failure of the cooling apparatus can lead to erratic behavior and—uncorrected—can damage the chip. If a customer complains of strange noises inside the PC, the CPU fan is a good place to look. As their bearings age, they start to whine.

Pentium

The Pentium (Series I) offers the following features:

- Speeds of 60 to over 200 MHz.
- 32-bit address bus and 32-bit registers.
- 64-bit data path to improve the speed of data transfers.
- Dual pipeline, 32-bit data bus that allows the chip to process two separate lines of code simultaneously.
- At least 8-KB write-back cache for data and an 8-KB write-through cache for programs.
- "Branch prediction"—in which the program cache attempts to anticipate branching within the code. The CPU stores a few lines of code from each branch so that when the program reaches the branch, the Pentium already has the code stored within the cache.

The following table lists the first generation of Pentium and Pentium-compatible chips.

Chip	Speed (MHz)	Register Width	External Data Bus	Address Bus	Internal Cache
Intel Pentium	60, 66	32-bit	64-bit	32-bit	8 KB W/B and 8 KB W/T
Intel Pentium	75	32-bit	64-bit	32-bit	8 KB W/B and 8 KB W/T
Intel Pentium	90, 100	32-bit	64-bit	32-bit	8 KB W/B and 8 KB W/T
Intel Pentium	120, 130	32-bit	64-bit	32-bit	8 KB W/B and 8 KB W/T
Intel Pentium	150, 166	32-bit	64-bit	32-bit	8 KB W/B and 8 KB W/T
Intel Pentium	180, 200	32-bit	64-bit	32-bit	8 KB W/B and

					8 KB W/T
CYRIX 6x86 (P-rating)	100, 120, 133, 200	32-bit	64-bit	32-bit	8 KB W/B and 8 KB W/T
AMD K5 (P-rating)	75, 90	32-bit	64-bit	32-bit	8 KB W/B and 8 KB W/T

Superscalar Technology

The main components of a processor—registers, decoders, and ALUs (arithmetic/ logic units) are collectively known as the *instruction pipeline*. To carry out a single instruction, a processor must:

- Read the instruction.
- Decode the instruction.
- Fetch operands (for math functions).
- Execute the instruction.
- Write back the results.

When possible, the Pentium processor breaks up a program into discrete tasks that are then shared between the pipelines, thus allowing the Pentium to execute two simple instructions simultaneously. Software must be specifically written to take advantage of this innovative feature, which is known as *multithreading.*

Pentium On-Board Cache

The original Pentium series came with two 8-KB caches—one for data and one for program code, compared with the single 8-KB cache on the 486 (16 KB on the DX4). As described with the 486 chip, the cache uses a technique called "branch prediction" to improve its ability to guess what data or program code will be required next by the processor.

Intel's Competitors

Competitors have moved away from simply making clones of the Intel processors. They are currently designing their own processors with unique features:

- NextGen Nx586
- AMD AmSx86
- Cyrix 6x86
- IBM 6x86

Pentium Pro

Intel made CPU selection even more complex with the introduction of the Pentium Pro, offering varied features, in different models, of the Pentium design. This processor was aimed at a 32-bit server and workstation-level applications such as computer-aided design (CAD), mechanical engineering, and advanced scientific computation. The Pentium Pro was packaged with a second speed-enhancing cache memory chip, and boasted 5.5 million transistors.

The main CPU and 16-KB first-level (L1) cache consist of 5.5 million transistors; the second chip is a 256- or 512-KB second-level (L2) cache with 15 million transistors. A 133-MHz Pentium Pro processes data about twice as fast as a 100-MHz Pentium.

Pentium MMX

Soon, more choices were on the way. About the time the 166-MHz Pentiums shipped, Intel introduced MMX (multimedia extension) technology, designed to enhance performance of data-hungry applications like graphics and games. With larger data and code caches, Pentiums with MMX technology can run non-MMX-enhanced software approximately 10 to 20 percent faster than a non-MMX CPU with the same clock speed.

The MMX chips also take advantage of dynamic branch prediction using the *branch target buffer* (BTB) to predict the most likely set of instructions to be executed.

The MMX Pentium processor is also more compatible with older 16-bit software than is the Pentium Pro;

Pentium II

1997, multimedia was becoming mainstream, and high performance in a graphical user environment was critical to CPU market success. The first 233-MHz, 7.5 million-transistor, Pentium II processor incorporated MMX technology and was packaged with a high-speed cache memory chip. Intel released Pentium II versions operating at speeds of up to 450 MHz.

The Pentium II incorporated the features of its older designs and added a number of enhancements. Among these are:

- Multiple Branch Prediction: predicts program execution through several branches, accelerating the flow of work to the processor.
- Data-flow Analysis: Creates an optimized, reordered schedule of instructions by analyzing data dependencies between instructions.

- Speculative Execution: Carries out instructions speculatively and, based on this optimized schedule, ensures that the processor's superscalar execution units remain busy, boosting overall performance.
- Single-edge connector (SEC) cartridge packaging: Developed by Intel, this enables high-volume availability and offers improved handling protection and a common form factor for future high-performance processors. This development resolved problems caused when pins were accidentally bent during installation or removal of CPUs.
- High-performance Dual Independent Bus (DIB) architecture (system bus and cache bus).
- System bus that supports multiple outstanding transactions to increase bandwidth availability. It also provides "glueless" support for up to two processors. This enables low-cost, two-way symmetric multiprocessing, providing a significant performance boost for multitasking operating systems and multithreaded applications. Many inexpensive motherboards offer two Slot 1 sockets, making it easy to build a dual processor system for use with operating systems like Windows NT or 2000.
- 512-KB unified, nonblocking, L2 cache: Improves performance by reducing average memory access time and providing fast access to recently used instructions and data. Performance is enhanced through a dedicated 64-bit cache bus. The speed of the L2 cache scales with the processor core frequency. This processor also incorporates separate 16-KB, L1 caches: one for instructions and one for data.
- Models available in 450, 400, and 350 MHz: Support memory caches for up to 4 GB of addressable memory space.
- Error correction coding (ECC) functionality on the L2 cache bus: for applications in which data intensity and reliability are essential.
- Pipelined floating-point unit (FPU): supports the 32-bit and 64-bit formats specified in IEEE (Institute of Electrical and Electronics Engineers) standard 754, as well as an 80-bit format.
- Parity-protected address/request and response system bus signals, with a retry mechanism for high data integrity and reliability.

Pentium III Processor

The Intel Pentium III processor is the member of the P6 family. With 28 million transistors, speeds from 500 to 733 MHz, and system bus speeds of 100 to 133 MHz, they mark a significant jump in PC CPU technology. They employ the same dynamic execution microarchitecture as the PII—a combination of multiple branch prediction, data-flow analysis, and speculative execution. This provides improved performance over older Pentium designs, while maintaining binary compatibility with all previous Intel processors. The Pentium III processor, shown in the picture below, also incorporates

MMX technology, plus streaming SIMD extensions for enhanced floating-point and 3-D application performance.



The Intel Pentium III Processor

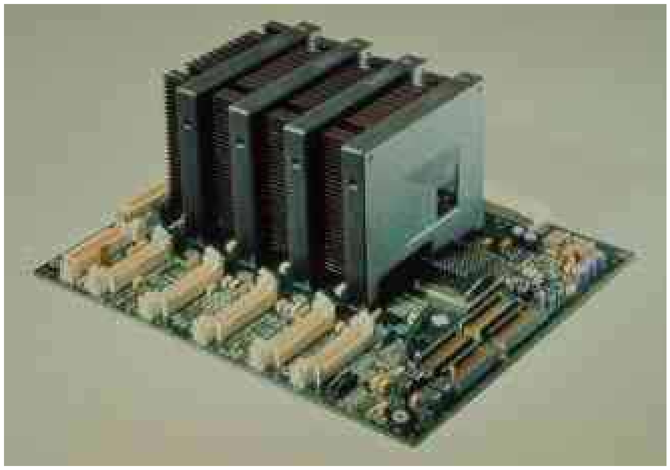
Intel offers a Xenon version of the Pentium III processor at 550 MHz, aimed at high-performance workstations and servers.

Pentium II Xeon

Xeon is a line of CPUs for use in servers. On July 26th, 1998 Intel introduced the Pentium II cartridge named Xeon. Aimed at servers and perhaps high-end users.

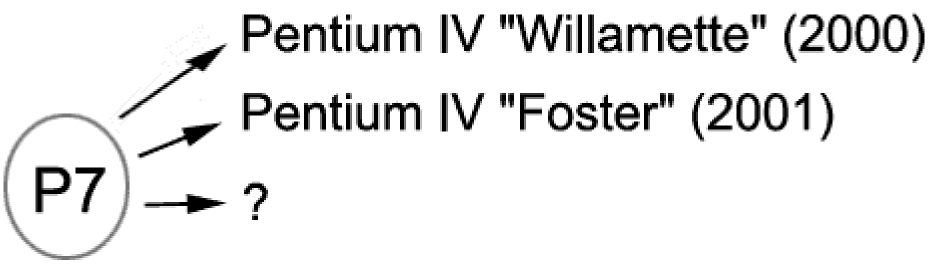
The Xeon is a Pentium II in a new cartridge fitting into a new connector called Slot Two. The module is twice as tall as the current Pentium II, but there are other important innovations and improvements:

- L2 cache RAM chips: CSRAM (Custom SRAM), which runs at full CPU speed.
- Different L2 cache sizes: 512, 1,024, or 2,048 KB L2 RAM.
- Up to 8 GB RAM can be cached.
- Up to four or even eight Xeons in one server.
- Support for clustered servers.
- New chip sets 82440GX and 82450NX.



The Pentium 4

The roadmap Intel made several years ago was to abandon the P6 core in favour of a completely new core. The processor codenamed "Willamette" should be the first of a new line of IA32 processors, which should be marketed side-by-side to the IA64 Itanium ("Merced"):



The Pentium IV

In November 2000 Intel introduced the new and very powerful high-end chip Pentium 4, formerly known as codename "Willamette". It was (as expected) delayed.

The Pentium 4 is a completely new processor holding several new designs. Here is a highlight:

- 400 MHz Front Side Bus of 128 bit width
- Execution Trace Cache
- 20 KB L1 cache and 256 KB L2
- The ALU (Arithmetical Logic Unit) runs at twice the clock speed
- A new socket for simple motherboard design
- Clock frequencies from 1500 MHz
- 20 stages pipeline
- SSE2 and 128 bit MMX
- 42 millions of transistors
- A new 423 pins socket design
- Dual Rambus memory channel with i850 chipset
- Only single processor mode available.

Heavy hardware

The first Pentium 4s were giant chips with a die size of 247 mm2. The 42 millions of transistors uses 60 watts and requires heavy cooling. However, Intel has done a lot to provide sufficient cooling using new materials and design.



The sockets

The first Pentium 4s came with a 423 pin design. These CPUs could only use RDRAM, which only requires few pins interface. Later came a 478 pin design with support for SDRAM.

The chipsets

The chip set i850 ("Tehama") is using a dual RDRAM bus. This is also heating up the MCH (north bridge) chip, so additional cooling is required.

The Intel i845 chipset was introduced August 2001. It gives an interface to standard 133 MHz SDRAM. This chipset is found in many cheap Pentium 4 computers.

Support for 266 MHz DDR RAM is found in the VIA P4X266 and P4X266A chipsets to be used with Pentium 4 (against Intels will; they don't like VIA at all).

SIS 645 is a chipset for Pentium 4 with support for both 266 and 333 MHz DDR RAM.

The i845D chipset

December 2001 Intel suddenly introduced a new flavour of the 845 chipset. It is designed to use with DDR-RAM!

There was no press releases about this product. Searching Intel's web for info on this chipset gave no result (December 27th 2001): However the i845D chipset was found in computers from Dell and others.

Intel has been under a hard pressure from AMD, VIA, and the Taiwanese motherboard companiess in all 2001, and it was fine to see, that they finally had to adapt to common sense. DDR is the RAM type to use - almost nobody wants RAMBUS! And Intel's dominating position in the market will promote better DDR RAM products.

Intel Processor Specifications

The processor can be thought of as the brains of the system and is responsible for executing software commands and performing calculation functions. The following table shows the features of the various Intel processors.

Chip Characteristics

Processor	Speed (MHz)	Heat Sink	Cooling Fan	Cache	Sockets	Pins
8088	5-8	No	No	No	DIP	40
80286	6 10 12	No	No	No	LLC PGA PLCC	68
80386SX	16 - 33	No	No	No	PGA	100
80386DX	16 - 33	No	No	No	PGA	100
80486SX	16 - 33	No	Yes on 33 MHz	0-256K	PGA	100
80486DX	25 - 50	No	Yes on 33 MHz	0-256k	PGA SQFP	168 208
Pentium	60-166	Yes	Yes	256-512k	PGA	296
Pentium Pro	233-266	Yes	Yes	256k-1mb	PGA	387
Pentium II	233-500	Yes	Yes	512k	SEC	242
Pentium III	450mhz-1.13ghz	Yes	Yes	256-512k	SEC/PGA	242/370

- With the Pentium MMX processors, 57 multimedia specific instructions were added to increase multimedia performance and increased the L1 cache size to 32KB.
- The Pentium Pro added Dynamic Execution and increase L2 cache to 512KB.
- The Pentium II had integrated MMX technology and used a new Single Edge Contact Cartridge(SEC).

- The Pentium III provided increased processor speeds, a 100mhz front size bus speed and increased L2 cache to 512KB.
- The Celeron processors are less expensive but only have a 66mhz bus and 128KB L2 cache.

While Intel holds the majority of the processor market share, companies such as AMD have been producing clones based on the X86 architecture. The table below outlines the various socket/slot types and the processors that they support.

Socket	Pins	Processor
Socket 4	237 PGA	Pentium 60/66, Pentium Overdrive
Socket 5	320 SPGA	Pentium 75-133, Pentium Overdrive
Socket 7	321 SPGA	Pentium 75-200, Pentium Overdrive
Socket 8	387 SPGA	Pentium Pro
Slot 1	242 SEC/SEPP	Pentium II, Pentium III, Celeron
Slot 2	330 SECC-2	Xeon
Super Socket 7	321 SPGA	Pentium MMX, Pentium Pro, AMD K6-2, K6-2+, K6-3, K6-3+
Socket 370	370 SPGA	Celeron, Pentium III, Cyrix III
Socket A	462 SPGA	AMD Athlon, Duron
Slot A	242 Slot A	AMD Athlon

Replacing and Upgrading Chips

Replacing a CPU can be very simple, but it is important to first carefully consider whether to do so. If you do decide to replace it, you will need to take care to avoid damaging the chip during installation. Before undertaking this process, always ask yourself, "What CPUs can be put on this motherboard?" The best source for an answer is the documentation packaged with the computer or motherboard. If the customer does not have the motherboard manual and you do not have a reference, the document should be available on the manufacturer's Web site.

Possible Upgrade Scenarios

There are a number of issues to think about when deciding whether to upgrade a CPU or replace a machine altogether. Perhaps the most important issue is the "value" of the upgrade. Will the suggested upgrade meet the operational requirements for that computer? There are limits to what can be upgraded and the results that can be expected from the upgrade. A poor upgrade can lead to total failure and, ultimately, require replacement of the motherboard. Again, the best source of information regarding CPU upgrades is the documentation that comes with the motherboard. The following table lists several possible scenarios for upgrading a CPU.

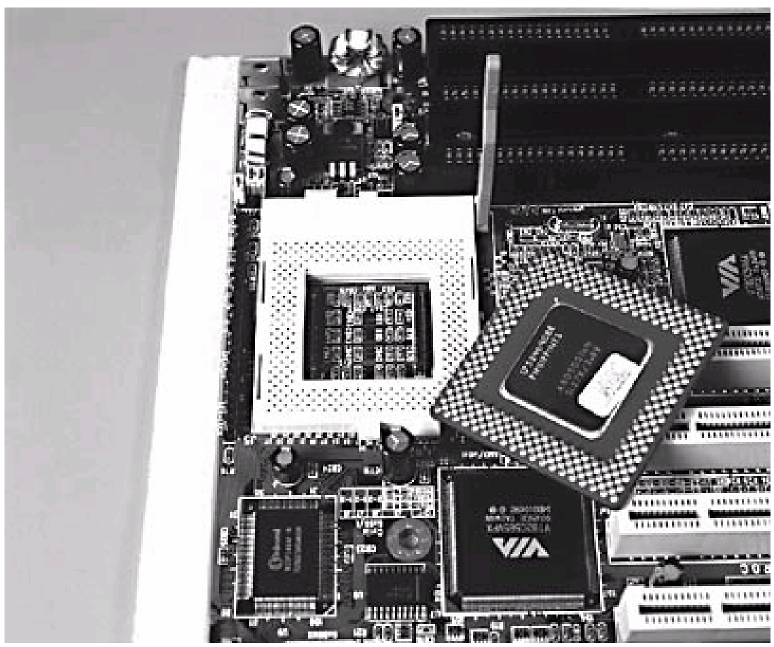
Existing CPU	Recommendation
8086/8088	Cannot be upgraded.
80286	Replace the motherboard with a new 486 or Pentium.
80386SX	Same as for 80286. Replace the motherboard.
80386DX	CPU has limited use—very slow. Upgrade to a Pentium motherboard.
486SX	If it has a good L2 cache (greater than 64 KB) and room to upgrade RAM, and VESA Local Bus (see Chapter 10, "Expansion Buses"), consider an AMD486DX/4 or a new Pentium motherboard.
486DX	Same as for 486SX.
486DX/2	Pentium upgrade chips will give you some improvement.
486DX/4	Same as for 486DX/2.
Pentium	Replace a Pentium 60, 66, or 75 with a faster CPU.

On average, it is more cost effective to replace an entire motherboard than it is to upgrade a CPU. However, you have to judge for yourself. Make sure that the new motherboard will fit into the computer case (check size and alignment of expansion buses) before starting the installation. Be sure that the power supply of the old case and new motherboard are of the same type with the proper connectors (such as AT, ATX).

Inserting a CPU

Several types of CPU sockets are sold. The three most common are:

- Low-insertion force (LIF).
- Zero-insertion force (ZIF).
- Single-edge connector (SEC).



ZIF socket with CPU inverted showing matching pins

A ZIF socket has a lever arm that allows for simple removal and installation of CPUs. ZIF socket had 169 pins and was used on 486SX systems.

The following table describes the types of ZIF sockets.

Upgrade Advice

Be very careful when handling a CPU or any exposed IC. Static discharge can damage or ruin the chip. Be sure to use a wrist-grounding strap or other approved antistatic device.

Take great care to not bend any pins, and make sure the CPU is properly lined up to seat Pin 1 by using the code notch.

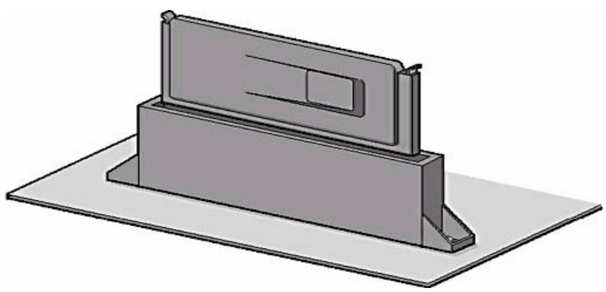
CAUTION

If you encounter any resistance, stop at once and figure out what is wrong.

Check the memory and bus speed required for the new CPU before attempting to boot the PC after the procedure. It might require new RAM and will most often demand that a jumper be set on the motherboard before operating at the new speed.

SEC Package/Slot 1 Upgrades

The Pentium II and III series are most commonly packaged in a SEC. This package, shown in the picture below, is very simple to work with. You will need a motherboard mount and might have to purchase a fan and heat sink before installing the CPU. Check the manual for jumper-setting adjustments, and follow the simple directions that come with the CPU.



Pentium Processor in an SEC Package and Slot 1

The actual task involves seating two plastic pins, sliding two guides over the sides of the CPU, then pushing the frame and CPU into the slot on the board. With dual CPU boards, you need to know which slot to use, and you might have to place a special card (which comes with the motherboard) in the second CPU position if it is to remain empty.

Chapter 5

Supplying Power to a Computer

OBJECTIVES

Power Supplies and Casing

Power Supplies Problems

Power Supply and Casing

It is a very important part and most often overlooked when the specifications of a computer are being written for a purchase or an upgrade. Computer owners boast of their motherboard, processor, RAM and Hard Disk etc. What they forget to care about is the power supply. In the following sections the power supply has been discussed in reasonable details as well as Surge suppressors, Phone-line surge protectors, Line conditioners, Uninterruptible power supplies (UPS) and RTC/NVRAM Batteries.

Primary Function and Operation

- The power supply provides the electrical power used by the computer.
- The power supply in a personal computer system is designed to convert 230V and 50Hz AC power into +3.3V, +5V, and +12V DC power.
- Some power supplies require you to switch between the two input ranges, whereas others auto-switch.

Positive DC Voltages

- Normally, the digital electronic components and circuits in the system (motherboard, adapter cards, and disk drive logic boards) use the +3.3V or +5V power, and the motors (disk drive motors and any fans) use the +12V power. The table below lists these devices and their power consumptions.

Power Consumption Ratings for PC Devices

Voltage	Devices Powered
+3.3V	Chipsets, DIMMs, PCI/AGP cards, miscellaneous chips
+5V	Disk drive logic, SIMMs, PCI/AGP cards, ISA cards, voltage regulators, miscellaneous chips
+12V	Motors, voltage regulators (high output)

- The power supply must deliver a good, steady supply of DC power so that the system can operate properly.
- Devices that run on voltages other than those specified in the above table must be powered by onboard voltage regulators. For example, RIMMs run on 2.5V that is supplied by an onboard regulator, and processors are supplied by a voltage regulator module (VRM) that normally is built into the motherboard as well.
- The main function of the +12V power is to run disk drive motors as well as the higher-output processor voltage regulators in some of the newer boards.

- Usually, a large amount of +12V current is available from the power supply, especially in those designed for systems with a large number of drive bays for disk drive motors.
- In the newer CPU voltage regulators, the +12V supply is used by any cooling fans in the system—which, of course, should always be running.
- A single cooling fan can draw between 100mA and 250mA (0.1–0.25 amps); however, most new fans use the lower 100mA figure.
- Most fans in desktop systems run on +12V, portable systems can use fans that run on +5V, or even +3.3V.

Negative DC Voltages

- PC power supply also generates–5V and –12V.
- The –5V is required on the Industry Standard Architecture (ISA) bus for full backward compatibility.
- –5V and –12V are supplied to the motherboard via the power supply connectors, the motherboard normally uses only the +3.3V, +5V, and +12V.
- The –5V is simply sent to the ISA bus on pin B5 so any ISA cards can use it.
- The analog data separator circuits found in older floppy controllers do use –5V.
- The motherboard logic normally doesn't use –12V, but it might be used in some board designs for serial port or LAN circuits.
- The older serial port circuits used +/–12V outputs, today most run on only +3.3V or +5V.

The Power_Good Signal

- The power supply actually prevents the computer from starting up or operating until all the power supply voltages are within the proper ranges.
- The power supply completes internal checks and tests before allowing the system to start. If the tests are successful, the power supply sends a special signal to the motherboard, called Power_Good.
- This signal must be continuously present for the system to run. Therefore, when the AC voltage dips and the power supply cannot maintain outputs within regulation tolerance, the Power_Good signal is withdrawn (goes low) and forces the system to reset. The system will not restart until the Power_Good signal returns.
- The Power_Good signal (sometimes called Power_OK or PWR_OK) is a +5V (nominal) active high signal (with variation from +2.4V through +6.0V generally being considered acceptable) that is supplied to the motherboard when the power supply has passed its internal self tests and the output voltages have stabilized.
- Some cheaper power supplies do not have proper Power_Good circuitry and might just tie any +5V line to that signal. Some motherboards are more sensitive to an improperly designed or improperly functioning Power_Good signal than others.

Intermittent startup problems are often the result of improper Power_Good signal timing.

- A common example is when you replace a motherboard in a system and then find that the system intermittently fails to start properly when you turn on the power.

Note

This can be very difficult to diagnose as the problem seems to be caused by the new motherboard. Although it seems as though the new motherboard is defective, it usually turns out that the power supply is poorly designed. It neither cannot produce stable enough power to properly operate the new board nor has an improperly wired or timed Power_Good signal (which is more likely). In these situations, replacing the supply with a higher quality unit, in addition to the new motherboard, is the proper solution.

Power Supply Form Factors

- The shape and general physical layout of a component is called the *form factor*.
- ATX and the related standards that followed have different connectors with additional voltages and signals that allow systems with greater power consumption and additional features that would otherwise not be possible with the AT style supplies.
- The power supply in a PC is described as a ***constant voltage half-bridge forward converting switching power supply***:
  - ***Constant voltage*** means that the power supply puts out the same voltage to the computer's internal components, no matter what the voltage of AC current running it or the capacity (wattage) of the power supply.
  - ***Half-bridge forward converting switching*** refers to the design and power regulation technique used by most suppliers. This design is commonly referred to as a *switching supply*. Compared to other types of power supplies, this design provides an efficient and inexpensive power source and generates a minimum amount of heat. It also maintains a small size and a low price.
- Plugging into the motherboard is one thing, but for the power supply to physically work in the system, it must fit the case. The bottom line is that it is up to you to make sure the power supply you purchase not only plugs into your motherboard but also fits into the chassis or case you plan to use.

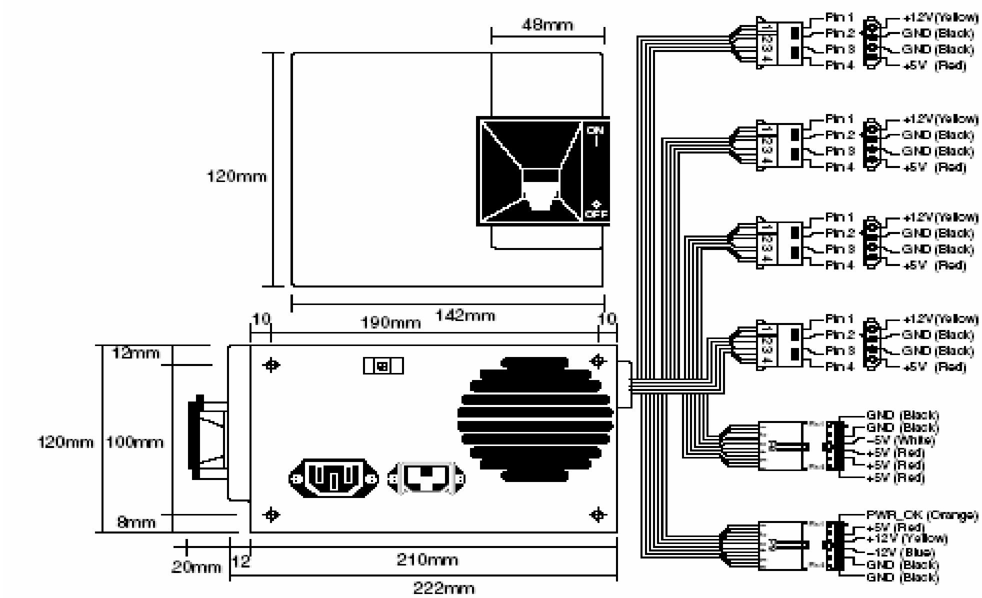
The table below shows the two most common power supply form factors, their connection types, and associated motherboards.

Power Supply Connector Types and Form Factors

Obsolete PS Form Factors	Originated From	Connector Type	Associated MB Form Factors

PC/XT style	IBM PC, PC-XT (1981/1983)	AT	PC/XT, Baby-AT
Modern PS Form Factors	**Originated From**	**Connector Type**	**Associated MB Form Factors**
ATX style	Intel ATX, ATX12V (1985/2000)	ATX	ATX, NLX, Micro-ATX

ATX has become by far the dominant form factor for power supplies, The PC/XT power supply and connectors are shown in the picture below. This form factor, however, is not used in modern systems today.

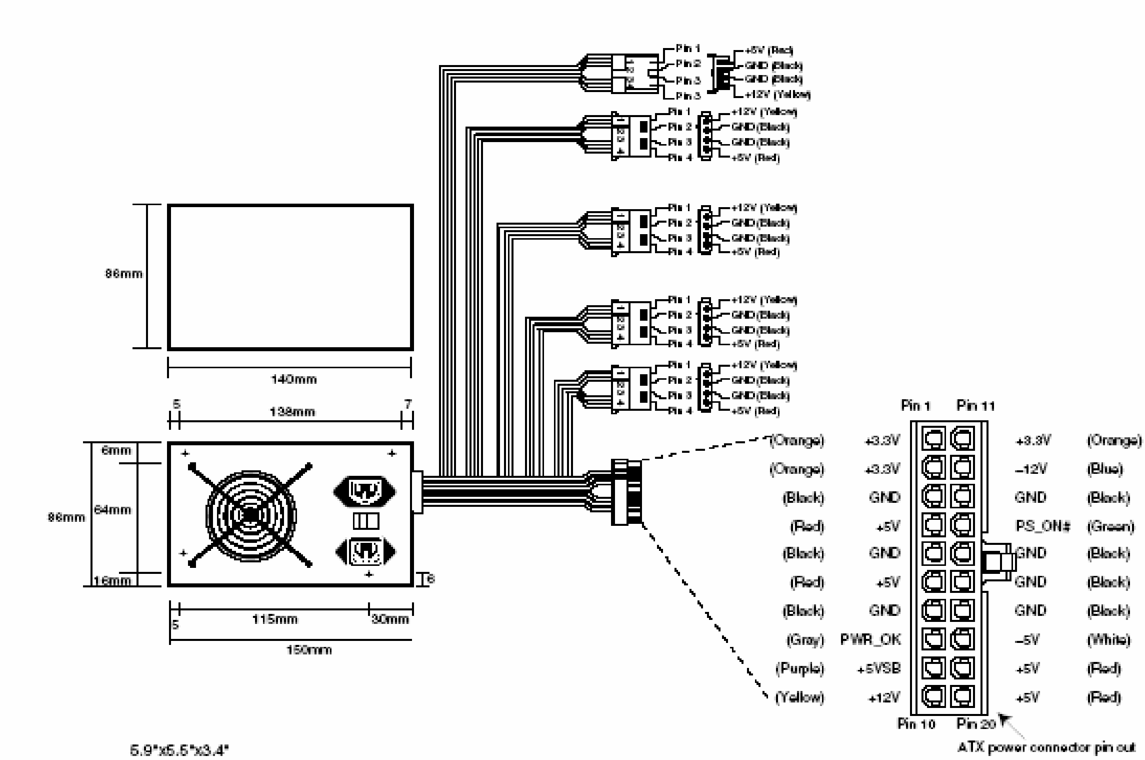


PC/XT Power Supply and Connectors

ATX Style

- In the ATX type power supply the fan should be mounted along the inner side of the supply drawing air in from the rear of the chassis and blow it inside across the motherboard. This kind of airflow runs in the opposite direction as most standard supplies, which exhaust air out the back of the supply through a hole in the case where the fan is mounted. The idea was that the reverse flow design could cool the system more efficiently with only a single fan, eliminating the need for a fan (active) heat sink on the CPU.

- Another benefit of the reverse-flow cooling is that the system would run cleaner, freer from dust and dirt. The case would be pressurized, so air would be continuously forced out of the cracks in the case.
- For systems that operate in extremely harsh environments, you can add a filter to the fan intake vent to further ensure that all the air entering the system is clean and free of dust.



ATX Power supply and Connectors

- Newer CPUs create more and more heat, the cooling capability of the system becomes more critical. It was found that using a standard negative pressure system with an exhaust fan on the power supply and an additional high-quality cooling fan blowing cool air right on the CPU is the best solution. For this reason, the ATX power supply specification has been amended to allow for either positive- or negative-pressure ventilation.
- Power supplies used with Baby-AT boards have two connectors that plug into the motherboard. If these connectors are inserted out of their normal sequence, the motherboard will burn out. Most responsible system manufacturers “key” the motherboard and power supply connectors so that you cannot install them backward or out of sequence. However, some vendors of cheaper systems do not feature this keying on the boards or supplies they use.
- The ATX form factor includes different power plugs for the motherboard to prevent users from plugging in their power supplies incorrectly. The ATX design features up to three motherboard power connectors that are definitively keyed, making plugging them in backward virtually impossible.

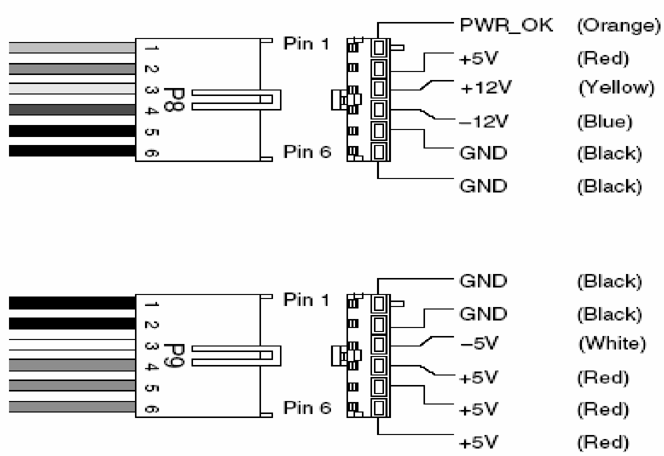
- The new ATX connectors also supply +3.3V, reducing the need for voltage regulators on the motherboard to power the chipset, DIMMs, and other +3.3V circuits.
- Besides the new +3.3V outputs, another set of outputs is furnished by an ATX power supply that is not normally seen on standard power supplies. The set consists of the Power_On (PS_ON) and 5V_Standby (5VSB) outputs mentioned earlier, known collectively as *Soft Power*. This enables features to be implemented, such as Wake on Ring or Wake on LAN, in which a signal from a modem or network adapter can actually cause a PC to wake up and power on. Many such systems also have the option of setting a wake-up time, at which the PC can automatically turn itself on to, perform scheduled tasks. These signals also can enable the optional use of the keyboard to power the system on. Users can enable these features because the 5V Standby power is always active, giving the motherboard a limited source of power even when off.

Motherboard Power Connectors

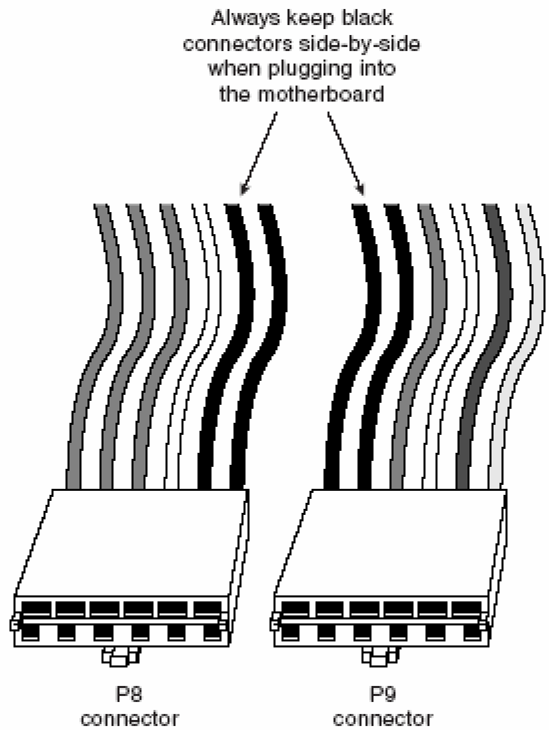
Every PC power supply has special connectors that attach to the motherboard, giving power to the system processor, memory, and all slotted add-on boards (ISA, PCI, and AGP). Attaching these connectors improperly can have a devastating effect on your PC, including burning up both your power supply and motherboard.

AT Power Supply Connectors

Industry standard PC, XT and AT, motherboards all use the same type of main power supply connectors. These supplies feature two main power connectors (P8 and P9), each with 6 pins that attach the power supply to the motherboard. Those two connectors are shown in the picture below.



AT main P8/P9 (sometimes also called P1/P2) power connectors.



The P8/P9 power connectors (sometimes also called P1/P2) that connect an AT power supply to the motherboard.

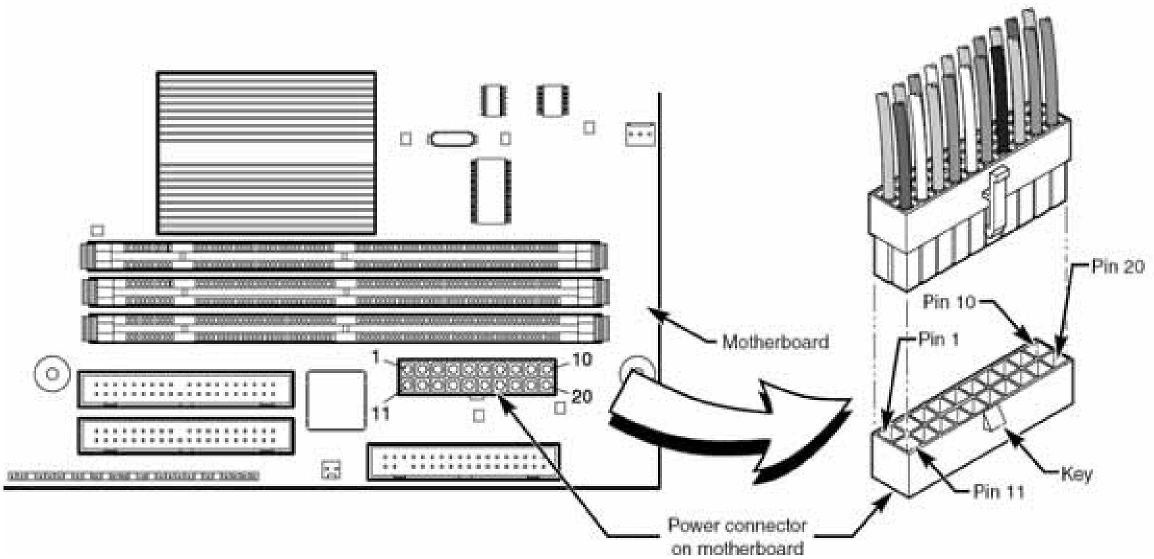
AT Power Supply Connectors

Connector	Pin	Signal	Color	Connector	Pin	Signal	Color
P8 (or P1)	1	Power_Good (+5V)	Orange	P9 (or P2)	1	Ground	Black
	2	+5V*	Red		2	Ground	Black
	3	+12V	Yellow		3	–5V	White
	4	–12V	Blue		4	+5V	Red
	5	Ground	Black		5	+5V	Red
	6	Ground	Black		6	+5V	Red

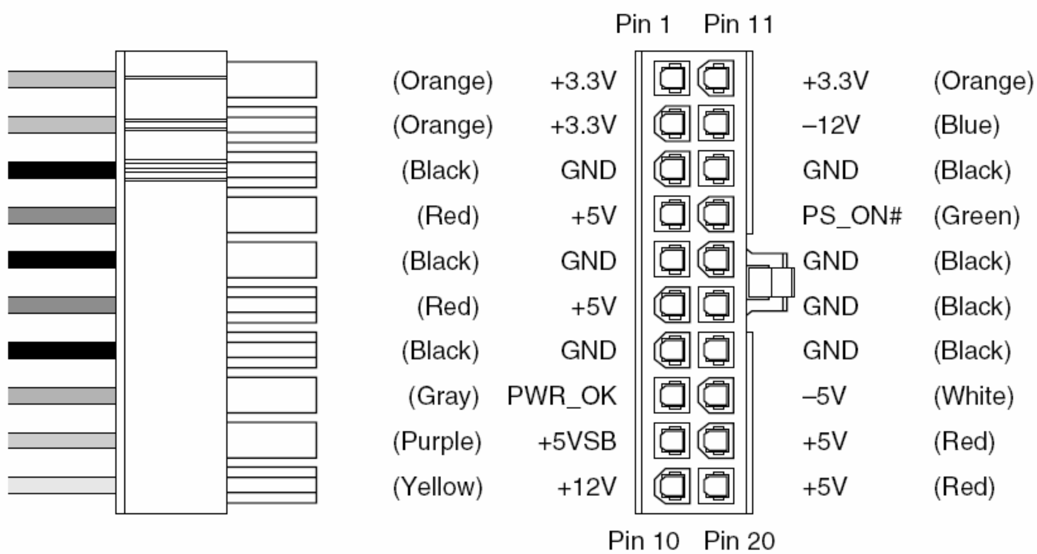
ATX Main Power Connector

The industry standard ATX power-supply–to–motherboard main connector is the Molex 39-29-9202 (or equivalent) 20-pin ATX style connector (see Figure 2.5). First used in the ATX form factor power supply, it also is used in other ATX-based variations.

This is a 20-pin keyed connector with pins configured as shown in the table. The colors for the wires listed are those recommended by the ATX standard; however, they are not required for compliance to the specification, so they could vary from manufacturer to manufacturer.



ATX style 20-pin motherboard main power connector perspective view.



ATX 20-pin main power connector terminal side view.

ATX Main Power Supply Connector Pinout (Wire Side View)

Color	Signal	Pin	Pin	Signal	Color
Orange	+3.3V	11	1	+3.3V	Orange
Blue	–12V	12	2	+3.3V	Orange

Black	GND	13	3	GND	Black
Green	PS_On	14	4	+5V	Red
Black	GND	15	5	GND	Black
Black	GND	16	6	+5V	Red
Black	GND	17	7	GND	Black
White	–5V	18	8	Power_Good	Gray
Red	+5V	19	9	+5VSB (Standby)	Purple
Red	+5V	20	10	+12V	Yellow

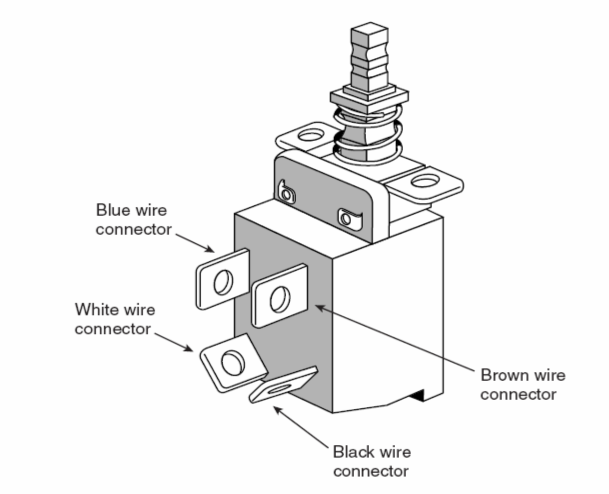
ATX12V Connector

Power for the processor comes from a device called the voltage regulator module (VRM), which is built into most modern motherboards. This device senses the CPU voltage requirements (usually via sense pins on the processor) and calibrates itself to provide the proper voltage to run the CPU. VRM is mostly designed to use 12V for input power.

Power Switch Connectors

The four or five wires are color-coded as follows:

- ***Brown and blue*** These wires are the live and neutral feed wires from the 220V power cord to the power supply. These are always hot when the power supply is plugged in.
- ***Black and white*** These wires carry the AC feed from the switch back to the power supply. These leads should be hot only when the power supply is plugged in and the switch is turned on.
- ***Green or green with a yellow stripe*** This is the ground lead. It should be connected to the PC case and should help ground the power supply to the case.



Power supply remote push button switch connections.

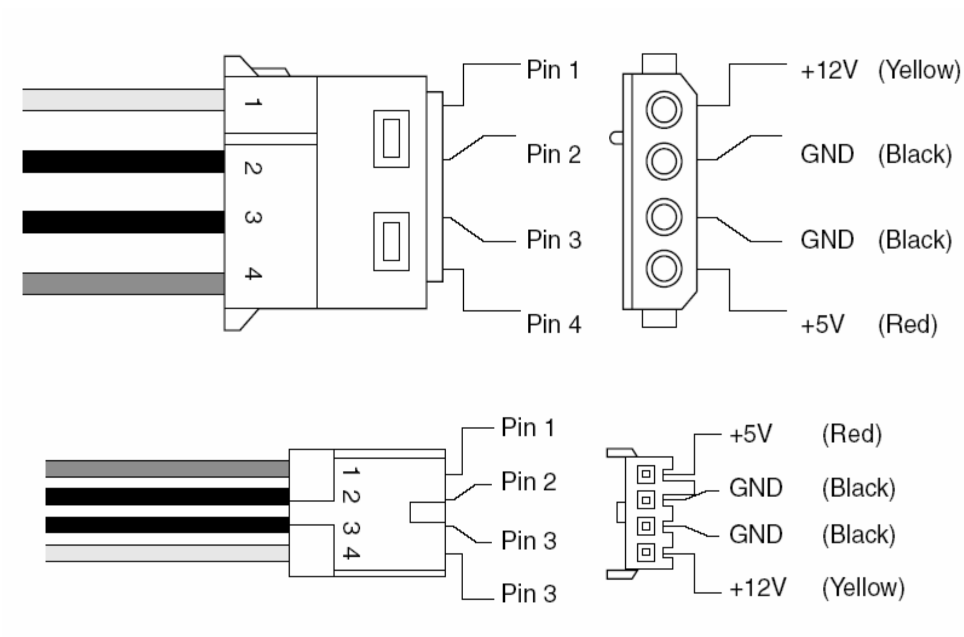
> **Warning**
> There is a continuous presence of the +5VSB (Standby) signal on pin 9 of the ATX connector means that the motherboard always is receiving standby power from the power supply when connected to an AC source, even when the computer is turned off. As a result, it is even more crucial to unplug an ATX system from the power source before working inside the case than it is on an earlier model system.

Peripheral Power Connectors

In addition to the motherboard power connectors, power supplies include a variety of peripheral power connectors for everything from floppy and hard drives to internal case fans. The following sections discuss the various types of connectors you're likely to find in your PC.

Peripheral and Floppy Drive Power Connectors

The disk drive connectors on power supplies are fairly universal with regard to pin configuration and even wire color. The picture below shows the peripheral and floppy power connectors.



Peripheral and floppy power connectors.

Peripheral Power Connector Pinout (Large Drive Power Connector)

Pin	Signal	Color	Pin	Signal	Color
1	+12V	Yellow	3	Gnd	Black

2	Gnd	Black	4	+5V	Red

Peripheral Power Connector Pinout (Small Drive Power Connector)

Pin	Signal	Color	Pin	Signal	Color
1	+5V	Red	3	Gnd	Black
2	Gnd	Black	4	+12	Yellow

Power Supply Loading

- PC power supplies are of a *switching* rather than a *linear* design. The switching type of design uses a high-speed oscillator circuit to convert the higher wall-socket AC voltage to the much lower DC voltage used to power the PC and PC components. Switching type power supplies are noted for being very efficient in size, weight, and energy in comparison to the linear design, which uses a large internal transformer to generate various outputs.
- One characteristic of all switching-type power supplies is that they do not run without a *load*. This means that you must have the supply plugged into something drawing power for the supply to work.
- If there is no load on the power supply on a bench, either the supply burns up or its protection circuitry shuts it down.
- High-quality switching power supplies have built-in load resistors and can run under a no-load situation because the supply loads.
- Other high-quality power supplies, such as those from PC Power and Cooling, have no internal load resistors. They require only a small load on the +5V line to operate properly. Many of the cheaper clone supplies, which often do not have built-in load resistors, might require +3.3V, +5V, and +12V loads to work.

Power Supply Ratings

The 300-watt and larger units are recommended for fully optioned desktops or tower systems. These supplies run any combination of motherboard and expansion card, as well as a large number of disk drives and other peripherals.

Power Supply Specifications

High-quality power supplies also help protect your systems. A power supply from a vendor such as Astec or PC Power and Cooling will not be damaged if any of the following conditions occur:

- A 100% power outage of any duration
- A brownout of any kind

- A spike of up to 2,500V applied directly to the AC input (for example, a lightning strike or a lightning simulation test)

Decent power supplies have an extremely low current leakage to ground of less than 500 microamps. This safety feature is important if your outlet has a missing or an improperly wired ground line.

Power Off When Not in Use

- Thermal expansion and contraction resulting from system start and stop remains the single largest cause of component failure.
- Thermal expansion and contraction affect not only chips and circuit boards, but also things such as hard disk drives.
- In essence, anything you can do to keep the system at a constant temperature
- Prolongs the life of the system, and the best way to accomplish this is to leave the system either permanently on or permanently off.
- It is recommended that the systems are powered on at the beginning of the workday and off at the end of the workday. Do not power the systems off for lunch, breaks, or any other short periods of time.
- Home users should leave their computer on, if it is going to be used later in the day or if instant access is important. They should turn off the system when leaving the house or when sleeping.
- Servers, of course, should be left on continuously.
- This seems to be the best compromise of system longevity with pure economics. No matter what, these are just guidelines; if it works better for you to leave your system on 24 hours a day, seven days a week, make it so.

Power Management

- As the standard PC configuration has grown to include capabilities formerly considered options, the power requirements of the system have increased steadily. Larger displays, CD-ROM drives, and audio adapters all need more power to run, and the cost of operating a PC rises steadily. To address these concerns, several programs and standards are now being developed that are intended to reduce the power needed to run a PC as much as possible.
- For standard desktop systems, power management is a matter of economy and convenience. By turning off specific components of the PC when they are not in use, you can reduce the electric bill and avoid having to power the computer up and down manually.
- For portable systems, power management is far more important. Adding CD-ROMs, speakers, and other components to a laptop or notebook computer reduces even further what is in many cases a short battery life. By adding new power

management technology, a portable system can supply power only to the components it actually needs to run, thus extending the life of the battery charge.

Energy Star Systems

The EPA (Environmental Protection Agency) has started a certification program for energy-efficient PCs and peripherals. To be a member of this program, the PC or display must drop to a power draw at the outlet of 30 watts or less during periods of inactivity. Systems that conform to this specification get to wear the Energy Star logo. This is a voluntary program; however, many PC manufacturers are finding that it helps them sell their systems if they can advertise these systems as energy efficient.
One problem with this type of system is that the motherboard and disk drives can go to sleep, which means they can enter a standby mode in which they draw very little power.

Advanced Power Management

Advanced Power Management (APM) is a specification jointly developed by Intel and Microsoft that defines a series of interfaces between power management–capable hardware and a computer's operating system. When it is fully activated, APM can automatically switch a computer between five states, depending on the system's current activity. Each state represents a further reduction in power use, accomplished by placing unused components into a low-power mode. The five system states are as follows:

- *Full On The* system is completely operational, with no power management occurring.
- *APM Enabled The* system is operational, with some devices being power managed. Unused devices can be powered down and the CPU clock slowed or stopped.
- *APM Standby The* system is not operational, with most devices in a low-power state. The CPU clock can be slowed or stopped, but operational parameters are retained in memory. When triggered by a specific user or system activities, the system can return to the APM Enabled state almost instantaneously.
- *APM Suspend The* system is not operational, with most devices unpowered. The CPU clock is stopped and operational parameters are saved to disk for later restoration. When triggered by a wakeup event, the system returns to the APM Enabled state relatively slowly.
- *Off* The system is not operational. The power supply is off.

APM requires support from both hardware and software to function. ATX-style power supplies can be controlled by software commands using the Power_On signal and the six-pin optional power connector.

Manufacturers are also integrating the same sort of control features into other system components, such as motherboards, monitors, and disk drives.

Operating systems that support APM, such as Windows 9x, trigger power management events by monitoring the activities performed by the computer user and the applications running on the system.

However, the OS does not directly address the power management capabilities of the hardware.

A system can have many hardware devices and many software functions participating in APM functions, which makes communication difficult. To address this problem, both the operating system and the hardware have an abstraction layer that facilitates communication between the various elements of the APM architecture. The operating system runs an APM driver that communicates with the various applications and software functions that trigger power management activities, while the system's APM-capable hardware devices all communicate with the system BIOS. The APM driver and the BIOS communicate directly, completing the link between the OS and the hardware. Thus, for APM to function, support for the standard must be built into the system's individual hardware devices, the system BIOS, and the operating system (which includes the APM driver). Without all these components, APM activities cannot occur.

Power Supply Troubleshooting

Troubleshooting the power supply basically means isolating the supply as the cause of problems within a system and, if necessary, replacing it.

Warning

It is rarely recommended that an inexperienced user open a power supply to make repairs because of the dangerous high voltages present. Even when unplugged, power supplies can retain dangerous voltage and must be discharged (like a monitor) before service.

Many symptoms indicate that the power supply in a system is failing.

In many cases a parity check error message can indicate a problem with the power supply. This might seem strange because the parity check message specifically refers to memory that has failed. The connection is that the power supply powers the memory, and memory with inadequate power fails. It takes some experience to know when this type of failure is power related and not caused by the memory. One clue is the repeatability of the problem. If the parity check message (or other problem) appears frequently and identifies the same memory location each time, it would be indicated that the defective memory is the problem. However, if the problem seems random, or if the memory location the error message points to as having failed seems random, it would indicate the defective power supply. The following is a list of PC problems that often are related to the power supply:

- Any power-on or system startup failures or lockups.
- Spontaneous rebooting or intermittent lockups during normal operation.
- Intermittent parity check or other memory-type errors.
- Hard disk and fan simultaneously failing to spin (no +12V).
- Overheating due to fan failure.
- Small brownouts cause the system to reset.
- Electric shocks felt on the system case or connectors.
- Slight static discharges disrupt system operation.

Irregular system problem indicate that the supply is defective.
Following symptoms point right to the power supply as a possible cause:

- System is completely dead (no fan, no cursor)
- Smoke
- Blown circuit breakers

If you suspect a power supply problem, some of the simple measurements and the more sophisticated tests outlined in this section can help you determine whether the power supply is at fault. Because these measurements might not detect some intermittent failures, you might have to use a spare power supply for a long-term evaluation. If the symptoms and problems disappear when a known good spare unit is installed, you have found the source of your problem.

Following points can help to zero in on common power supply related problems: -
1. Check AC power input. Make sure the cord is firmly seated in the wall socket and in the power supply socket. Try a different cord.
2. Check DC power connections. Make sure the motherboard and disk drive power connectors are firmly seated and making good contact. Check for loose screws.
3. Check DC power output. Use a digital multimeter to check for proper voltages. If it's below spec, replace the power supply.
4. Check installed peripherals. Remove all boards and drives and retest the system. If it works, add back in items one at a time until the system fails again. The last item added before the failure returns is likely defective.
Many types of symptoms can indicate problems with the power supply. Because the power supply literally powers everything else in the system, everything from disk drive problems to memory problems to motherboard problems can often be traced back to the power supply as the root cause.

Overloaded Power Supplies

The wattage rating can sometimes be very misleading. Not all 300-watt supplies are created the same. Many of the cheaper supplies provide noisy or unstable power, which can cause numerous problems with the system. Another problem with under-engineered power supplies is that they can run hot and force the system to do so as well. The

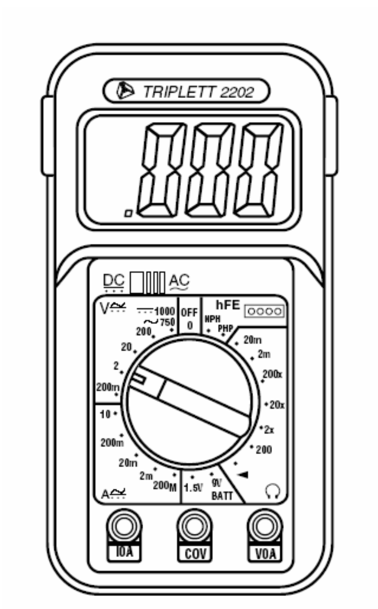
repeated heating and cooling of solid-state components eventually causes a computer system to fail. It is recommended that the original supply in a system be replaced with a heavier-duty model, which solves the problem. Because power supplies come in common form factors, finding a heavy-duty replacement for most systems is easy.

Using Digital Multimeters

One simple test you can perform on a power supply is to check the output voltages. This shows whether a power supply is operating correctly and whether the output voltages are within the correct tolerance range. Note that you must measure all voltages with the power supply connected to a proper load, which usually means testing while the power supply is still installed in the system and connected to the motherboard and peripheral devices.

Selecting a Meter

You need a simple Digital Multimeter (DMM) or Digital Volt-Ohm Meter (DVOM) to perform voltage and resistance checks on electronic circuits. You should use only a DMM instead of the older needle-type multimeters because the older meters work by injecting 9V into the circuit when measuring resistance, which damages most computer circuits.

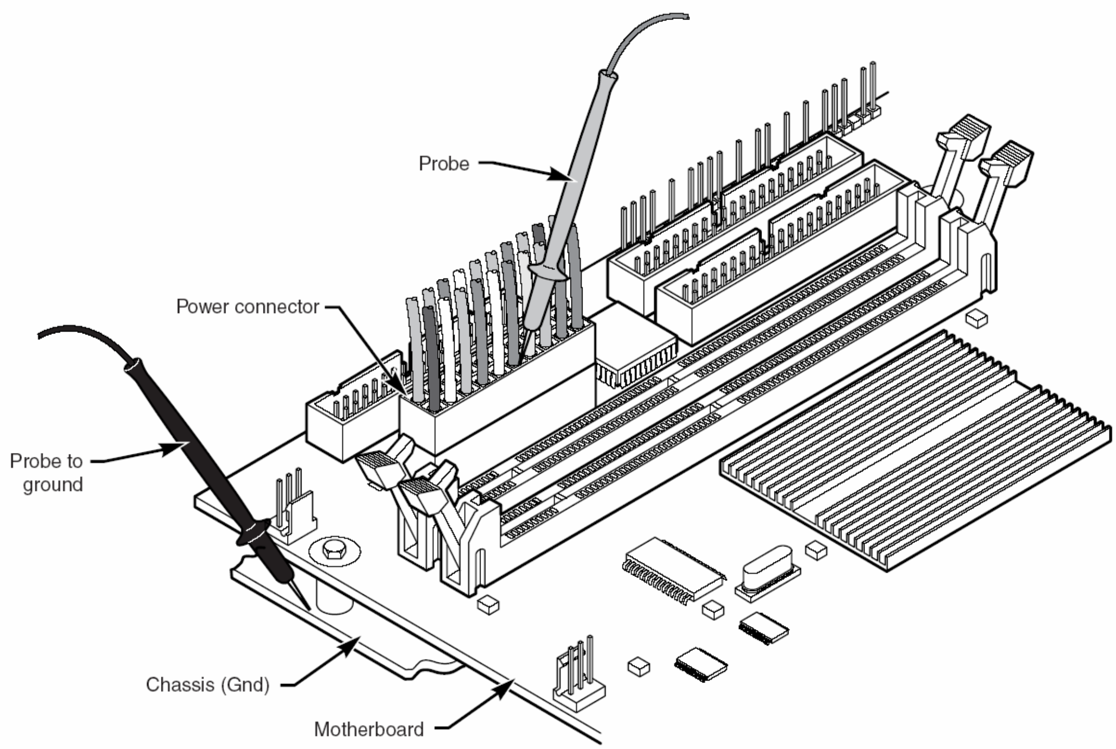


A typical DMM.

A DMM uses a much smaller voltage (usually 1.5V) when making resistance measurements, which is safe for electronic equipment. You can get a good DMM with many features from several sources.

Measuring Voltage

To measure voltages on a system that is operating, you must use a technique called *back probing* on the connectors (see Figure 5.10). You cannot disconnect any of the connectors while the system is running, so you must measure with everything connected. Nearly all the connectors you need to probe have openings in the back where the wires enter the connector. The meter probes are narrow enough to fit into the connector alongside the wire and make contact with the metal terminal inside. The technique is called back probing because you are probing the connector from the back. You must use this back-probing technique to perform virtually all the following measurements. To test a power supply for proper output, check the voltage at the Power_Good pin (P8-1 on AT, pin 8 on the ATX-type connector) for +3V to +6V of power. If the measurement is not within this range, the system never sees the Power_Good signal and therefore does not start or run properly. In most cases, the power supply is bad and must be replaced.



Back probing the power supply connectors.

Continue by measuring the voltage ranges of the pins on the motherboard and drive power connectors. If you are measuring voltages for testing purposes, any reading within 10% of the specified voltage is considered acceptable, although most manufacturers of high-quality power supplies specify a tighter 5% tolerance. For ATX power supplies, the specification requires that voltages must be within 5% of the rating, except for the 3.3V

current, which must be within 4%. The following table shows the voltage ranges within these tolerances.

Voltage ranges for an ATX power supplies

	Loose Tolerance		Tight Tolerance	
Desired Voltage	**Min. (10%)**	**Max. (+8%)**	**Min. (5%)**	**Max. (+5%)**
+3.3V	2.97V	3.63V	3.135	3.465
+/–5.0V	4.5V	5.4V	4.75	5.25
+/–12.0V	10.8V	12.9V	11.4	12.6

The Power_Good signal has tolerances that are different from the other voltages, although it is nominally +5V in most systems. The trigger point for Power_Good is about +2.4V, but most systems require the signal voltage to be within the tolerances listed here:

Voltage Ranges for the Power_Good Signal

Signal	Minimum	Maximum
Power_Good (+5V)	3.0V	6.0V

Replace the power supply if the voltages you measure are out of these ranges.
Again, it is worth noting that any and all power supply tests and measurements must be made with the power supply properly loaded, which usually means it must be installed in a system and the system must be running.

Power Protection Systems

The following systems would be explained here.

- Surge suppressors
- Phone-line surge protectors
- Line conditioners
- Standby power supplies (SPS)
- Uninterruptible power supplies (UPS)

Surge Suppressors

The Surge suppressor can absorb the high-voltage transients produced by nearby lightning strikes and power equipment. Surge protectors use several devices, usually metal-oxide varistors (MOVs) that can clamp and shunt away all voltages above a certain

level. MOVs are designed to accept voltages as high as 6,000V and divert any power above 200V to ground.
MOVs can handle normal surges, but powerful surges such as a direct lightning strike can blow right through them. Some surge protectors have status lights that let you know when a surge large enough to blow the MOVs has occurred.

The surge suppressors should bear the Underwriters Laboratories sign and should have following two features:

- Conformance to the UL 1449 or 1950 standard
- A status light indicating protection is on

Phone Line Surge Protectors

In addition to protecting the power lines, it is critical to provide protection to your systems from any connected phone lines. If you are using a modem or fax board that is plugged into the phone system, any surges or spikes that travel through the phone line can damage your system. In many areas, the phone lines are especially susceptible to lightning strikes, which are the leading cause of fried modems and damage to the computer equipment attached to them. Some of the standard power line surge protectors include connectors for phone line protection as well.

Line Conditioners

A line conditioner filters the power; bridges voltage drops, suppresses high-voltage and current conditions, and generally acts as a buffer between the power line and the system. A line conditioner does the job of a surge suppressor, and much more. It is more of an active device, functioning continuously, rather than a passive device that activates only when a surge is present. A line conditioner provides true power conditioning and can handle numerous problems. It contains transformers, capacitors, and other circuitry that can temporarily bridge a low-voltage situation.

Uninterruptible Power Supplies

An Uninterruptible Power Supply provides a power source that is conditioned and that cannot be interrupted. UPS is an online system because they continuously function and supply power to your computer systems. In a true UPS system the computer is always operating from the battery, there is no switching circuit. A voltage inverter converts from 12V DC to 110V AC. It is essentially a power system that generates power independently of the AC line. A battery charger connected to the line or wall current keeps the battery charged at a rate equal to or greater than the rate at which power is consumed.

When the AC current supplying the battery charger fails, a true UPS continues functioning undisturbed because the battery-charging function is all that is lost. Because the computer was already running off the battery, no switching takes place, and no power disruption is possible. The battery begins discharging at a rate dictated by the amount of load the system places on the unit gives you plenty of time to execute an orderly system shutdown.

Note
Many UPSes today come equipped with a cable and software that enables the protected computer to shut down in an orderly manner on receipt of a signal from the UPS. This way, the system can shut down properly even if the computer is unattended. Some operating systems, such as Windows NT/2000, contain their own UPS software components.

Warning
Do not connect a laser printer to a socket in any UPS unit unless it is designed for this purpose.

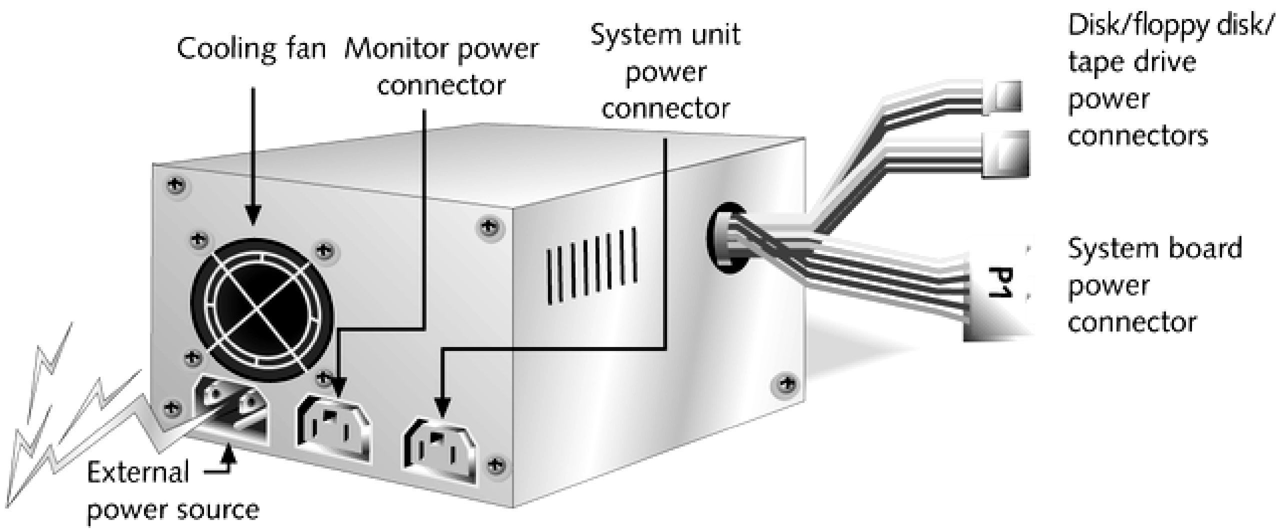
RTC/NVRAM Batteries

All 16-bit and higher systems have a special type of chip in them that combines a real-time clock (RTC) with at least 64 bytes (including the clock data) of Non-Volatile RAM (NVRAM) memory. This chip is officially called the RTC/NVRAM chip but is often referred to as the CMOS chip or CMOS RAM because the type of chip used is produced using a CMOS (Complementary Metal-Oxide Semiconductor) process. CMOS design chips are known for very low power consumption. This special RTC/NVRAM chip is designed to run off a battery for several years. These chips include a real-time clock. Its function should be obvious: The clock enables software to read the date and time and preserves the date and time data even when the system is powered off or unplugged. The NVRAM portion of the chip has another function. It is designed to store basic system configuration, including the amount of memory installed, types of floppy and hard disk drives, and other information. Some of the more modern motherboards use extended NVRAM chips with as much as 2KB or more of space to hold this configuration information. This is especially true for Plug-and-Play systems, which store not only the motherboard configuration but also the configuration of adapter cards. This system can then read this information every time you power on the system.

These chips normally are powered by some type of battery while the system is off. This battery preserves the information in the NVRAM and powers the clock. Most systems use a lithium-type battery because they have a very long life, especially at the low power draw from the typical RTC/NVRAM chip. Under normal conditions, the battery will last for 10 years. Symptoms that indicate that the battery is about to fail include having to

reset the clock on your PC every time you shut down the system (especially after moving it) and problems during the system's POST, such as drive-detection difficulties.

Warning
When you replace a PC battery, be sure you get the polarity correct; otherwise, you will damage the RTC/NVRAM (CMOS) chip. Because the chip is soldered onto most motherboards, this can be an expensive mistake!
When you replace a battery, in most cases the existing data stored in the NVRAM is lost. It is recommended that all the system configuration settings by running the BIOS Setup program and copy or print out all the screens showing the various settings. Some Setup programs offer the capability to save the NVRAM data to a file for later restoration if necessary.



Overview of Power Supplies

A standard power supply draws power from a local, alternating current (AC) source (usually a wall outlet) and converts it to either 3.3 or 5 volts direct current (DC), for on-board electronics, and 12 volts DC for motors and hard drives. In all cases, it delivers both positive and negative DC to the computer. Power supplies must "condition" the power, smoothing out any radical changes in its quality. Many homes and offices have power that fluctuates far more than the delicate parts of a PC can tolerate and survive. Most PC power supplies also provide the system's cooling and processor fans that keep the machine from overheating.

If the computer's power supply is providing reliable, clean power and its own cooling fan works, all is well. If the power supply or its fan should fail or cause erratic behavior by the PC, the power supply must be replaced. (While it is possible to remove and replace a power-supply fan, the low cost of a power supply makes it more practical to replace the power supply itself.)

Many newer supplies have a universal input that will accept either 110 VAC (volts alternating current), 60 Hz (U.S. standard power), or 220 VAC, 50 Hz (European/Asian standard). When replacing a power supply, there are three things to consider: physical size, wattage, and connectors. This chapter covers the basics of power supplies.

Power-Supply Sizes

Power supplies are available in a few standard sizes and shapes. However, the names for power supplies are anything but standard. They are based on the types of case they will be used in and the types of motherboard connections they will support. This is because different styles of cases place items such as plug fittings, mounting screws, and fans in different places. A few years ago, a new type of motherboard cable power fitting began to appear on the market. The older models are known as AT-style, and the newer ones are known as ATX. We cover both in this chapter, because you will need to be able to work with either one. The ATX design simplifies the placement of connections, so there is little to worry about with case compatibility. The main issues to be aware of are how much wattage the PC needs to power its parts and how many connectors for peripherals are required. Generally speaking, older Pentium-based computers and all 486-based and earlier PCs used AT supplies; almost all Pentium II and later-based systems use ATX supplies. The ATX design is preferred for two reasons:

- The on/off power control circuit (not the button) on ATX boards is built into the motherboard. On AT PCs, it comes from the power supply.
- AT power supplies connect to the motherboard through a pair of six-wire connectors. ATX power supplies connect through a single 20-pin connector.

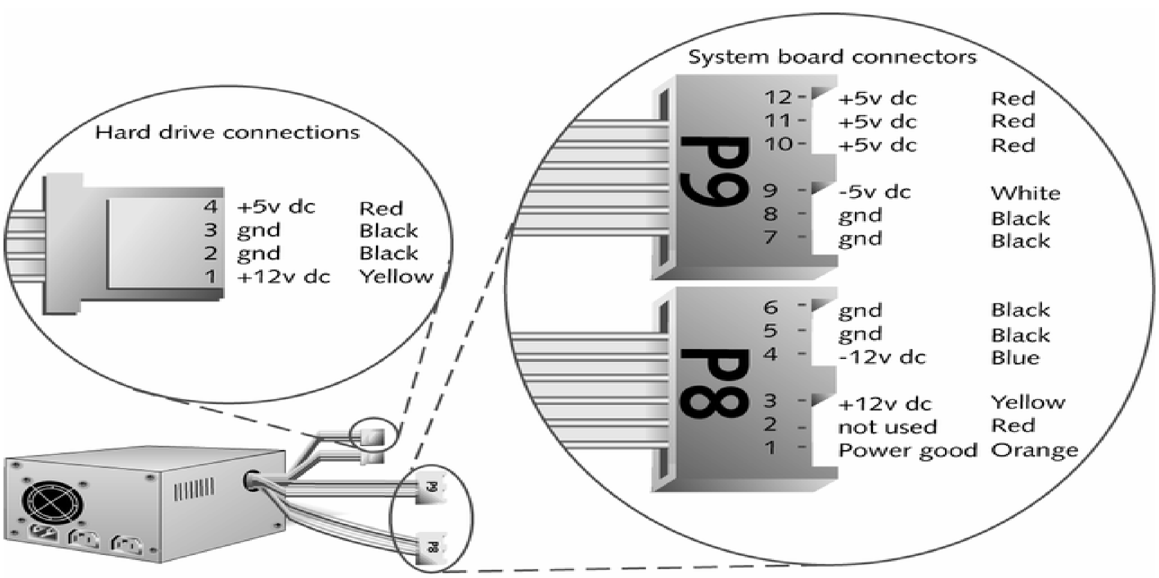
A few motherboards and power supplies provide both AT and ATX fittings and switch support. They are rare, but will open up more options if you have to repair such a system. Generally, you should use ATX for all replacements, if possible. It's a good idea to compare the existing power supply to the new one. Make sure that they are physically the same size, have the same connectors, and that the new one has at least the same power rating. Some high-quality power supplies offer "silencer" fans, that are much quieter than most models.

Power-Supply Wattage

Power supplies are rated according to the maximum sustained power (given in watts) that they can produce. A *watt* is a unit of electrical power equivalent to one volt-ampere. It is important to keep in mind that the power supply must produce at least enough energy to operate all the components of the system at one time. Don't rely on the computer's operating consumption alone, which can be obtained by adding the requirements for the devices in the PC. Remember that a much larger drain occurs as the machine powers up, when hard drives and other heavy feeders simultaneously compete for the available startup power. Most general-use computers require 130 watts while running and about 200-205 watts when booting. Sound cards, modems, and (worst of all) monitors attached with an accessory plug in the case can push a weak power supply to its limit and beyond. Servers and high-performance workstations often have an abundance of RAM, multiple drives, SCSI (Small Computer System Interface) cards, and power-hungry video adapters, along with one or more network cards. They often demand power supplies of 35-500 watts.

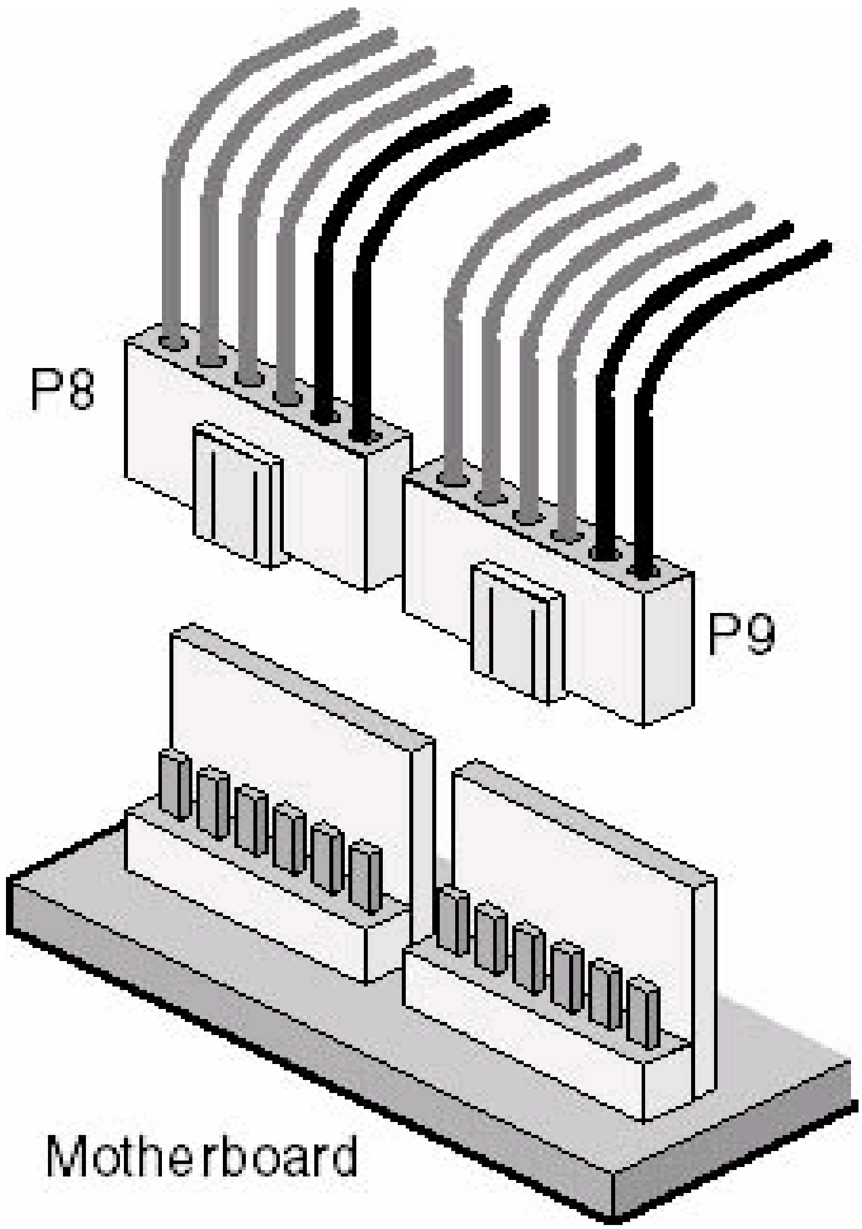
Power-Supply Connectors

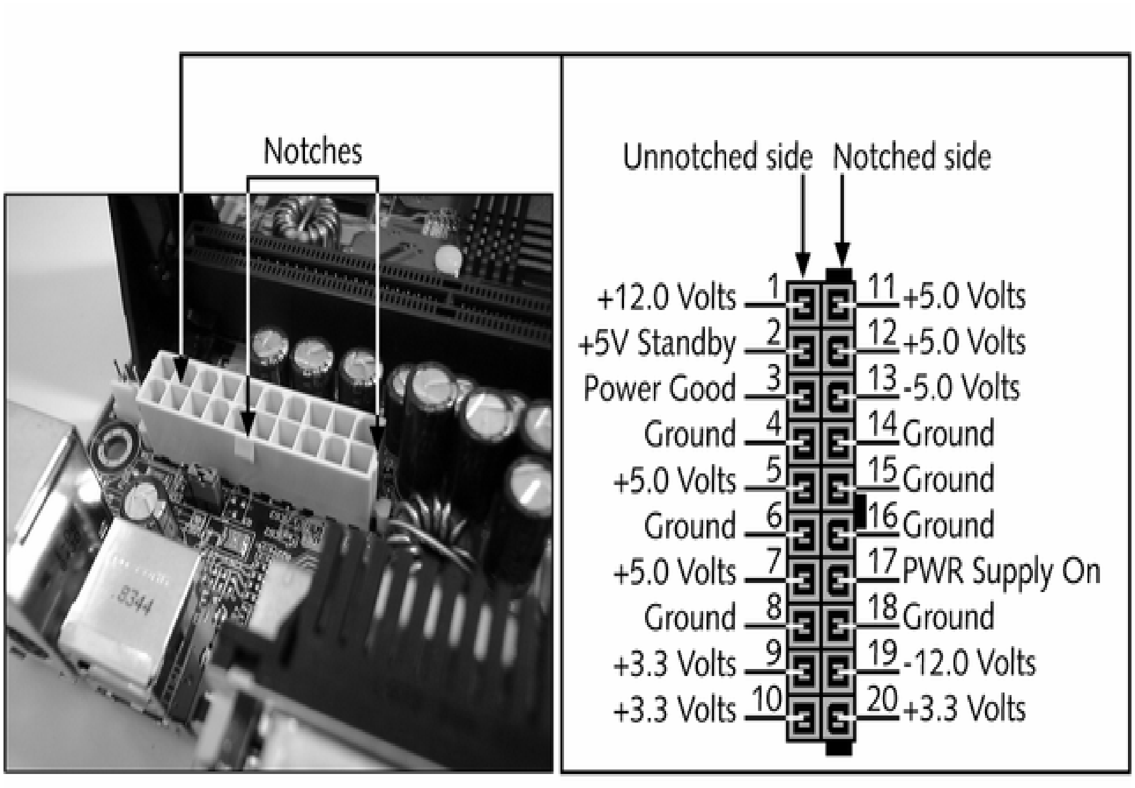
Power supplies employ several types of connectors; all are easy to identify and use. On the outside of the computer enclosure, a standard male AC plug and three-conductor wire (two power wires and a ground) draws current from a wall outlet, with a female connection entering the receptacle in the back of the power supply. On the inside are three types of connectors: the power main to the motherboard (which differ, as mentioned, in AT and ATX models) and two types of four-pin fittings to supply 5 volts and 3.3 volts of power to peripherals such as the floppy disk and hard disk drives. Let's take a close look at each in turn.



AT-Style Connections to the Motherboard

A pair of almost identical connectors, designated P8 and P9, links the power supply to the motherboard (see Figure 5.1). These connectors are seated into a row of six pins and matching plastic guides, or "teeth," on the motherboard. The P8 and P9 connectors *must* be placed in the proper orientation. The motherboard manual will show which fitting is for P8 and P9. If the connectors are not marked, make sure that the two black wires on each plug are side by side and that the orange wire (on P8) and the two red wires (on P9) are on the outside as you push them into place.





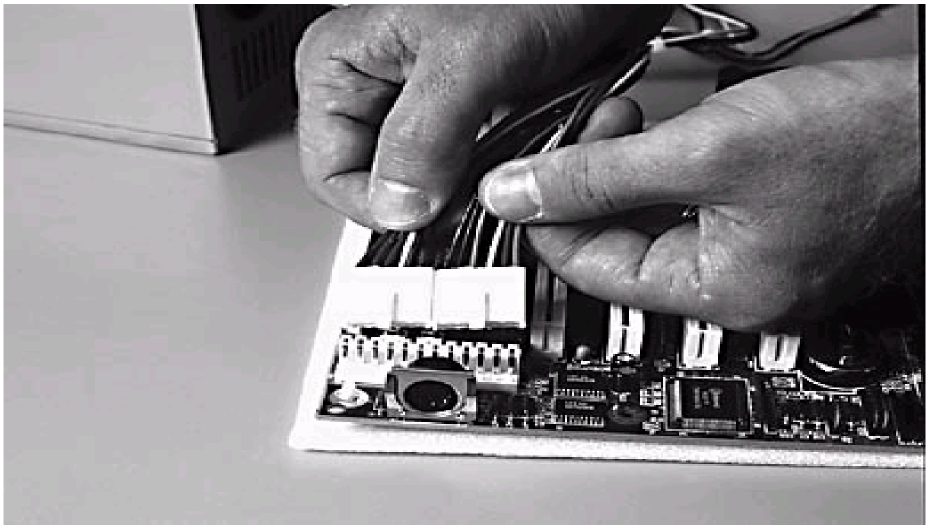
connectors and motherboard fitting

The following table of power cables shows voltage values for each of the color-coded wires on P8 and P9. The ground wires are considered 0 volts; all voltage measurements, "Basic Electricity") are taken between the black wires and one of the colored wires.

Cable Color	Supply In	Tolerance
Yellow	+12	±10%
Blue	-12	±10%
Red	+5	±5%
White	-5	±5%
Black	Ground	N/A

NOTE
Some computer makers employ proprietary power connections that require a special power supply. To install the new part, you will need to follow the instructions that are included with the computer.

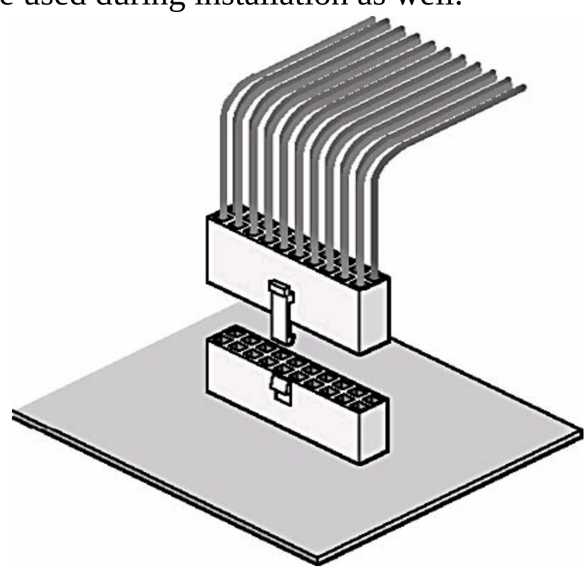
Remember to install the P8 and P9 plugs so that the black wires are side by side. Installing them on the wrong receptacle can damage both the motherboard and the power supply. The picture below shows the P8 and P9 connectors and a motherboard.



Connecting P8 and P9

ATX Motherboard Connections

The newer ATX main power connection is much easier to install. A single 20-wire plug is set into a fitted receptacle and secured with a catch on the side of the plug that snaps over the fitting. The picture below shows the parts being seated. A small, flat-tip screwdriver is a handy tool for easing the pressure on the catch to remove the plug. In some cases, it can be used during installation as well.

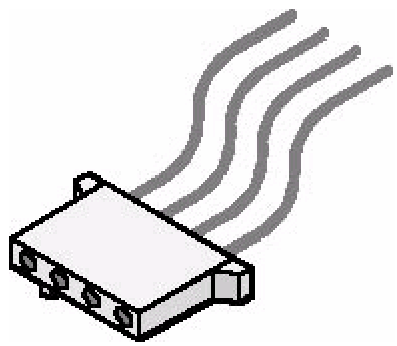


Placing an ATX plug in its motherboard receptacle

Connections to Peripheral Hardware

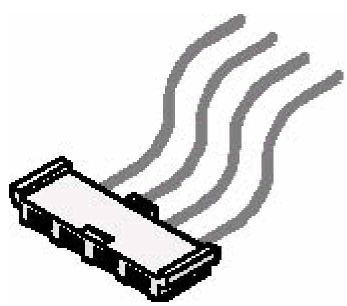
Two standard types of connectors can connect to peripheral hardware:

- **Molex connector:** This is the most commonly used power connector. It provides both 12-volt and 5-volt power. Hard disk drives, internal tape drives, CD-ROM drives, DVD (digital video disc) drives, and older 5.25-inch floppy disk drives all use this fitting. The Molex connector has two rounded corners and two sharp corners to ensure that it will be properly installed.



Molex connector (not to scale)

- **Mini connector:** Most power supplies provide one or more "mini" connectors. The mini, shown in the picture, is used primarily on 3.5-inch floppy-disk drives. It has four pin-outs and, usually, four wires. Most are fitted with keys that make it difficult, but not impossible, to install upside down. Be sure to orient the connector correctly; applying power with the connector reversed can damage or destroy the drive.



Mini connector (not to scale)

Two- and Three-Pin Mini Plugs

A less common type of power connector is used to connect the fan of a Pentium II or III processor to the motherboard for power, to connect a CD-ROM drive to a sound card,

and to provide power for 3.5-inch floppy disk drives. These connectors have two or three wires which are usually red and black or red, yellow, and black.

Power-Supply Problems

Power Failures

Power supplies are affected by the quality of the local power source. Common power delivery problems such as spikes, surges, sags, brownouts, and blackouts affect the stability and operation of the main power supply and are passed on to the computer. While most users don't notice sudden changes in the quality of electrical power, computers and other sensitive electronics do and while we can't fully control these problems, there are a few things we can do to protect our equipment and data and ensure a reasonably clean electrical supply.

Problem	Description
Surges	These are brief (and sometimes catastrophic) increases in the voltage source (very high voltage for a very short time). They can originate with the power source (the local power company), but most often are due to lightning strikes.
Spikes	Spikes are very short over voltage conditions. Spikes are measured in nanoseconds, while a surge is measured in milliseconds.
Sags	These are brief decreases of voltage at the power source.
Brownouts	If a sag lasts longer than one second, it is called a brownout. The overloading of a primary power source can cause brownouts. Some brownouts are "scheduled" by power companies to prevent overloading of circuits and potential catastrophic failure of the system.
Blackout	This is a complete power failure, which can be caused by equipment failure (local or regional) or accidental cutting of power cables. When the power returns after a blackout, there is a power spike and danger of a power surge.

Power-Protection Devices

Surge suppressors are devices used to filter out the effects of voltage spikes and surges that are present in commercial power sources and smooth out power variations. They are

available from local computer dealers and superstores. A good surge suppressor will protect your system from most problems, but if you purchase an economy model, it might not work when you need it most. Keep in mind that almost nothing will shield your hardware from a very close lightning strike.

A *UPS* is an inline battery backup. When properly installed between a computer and the wall outlet, a UPS device protects the computer from surges and acts as a battery when the power dips or fails. It also provides a warning that the power is out of specification (above or below acceptable levels). Many models can also interact with the computer and initiate a safe shutdown in the event of a complete power failure. They do this by means of software that runs in the background and is set in action by a signal through one of the computer's COM ports when the power goes down. The amount of time that a UPS device can keep a system running is determined by battery capacity and the power demands of the equipment connected to it. A more powerful UPS device will need its own line and circuit breaker.

One of the principal power drains is the monitor. To keep a system online as long as possible during a power failure, turn off the monitor immediately after the failure commences. When recommending a UPS, take into consideration how much protection is needed as well as the importance of peace of mind to the user. The VA rating (voltage x amps = watts) must be sufficient to supply the computer and all its peripherals with power for long enough to safely shut down the system. The easiest way to calculate this number is to total the power rating (watts) for *all* pieces of equipment that are to be connected to the UPS, as shown in the following table.

Device	Power Rating (Watts)	Connected to UPS	Power Required
Computer	200-350	Yes	250
Monitor	80-100	Yes	80
External modem	5.5	No	0
External backup drive	50	Yes	50
Total	330-500	—	380

Power-Supply Problems

The most easily recognized problem is a complete failure of the power supply. This is easy to detect because in the event of a failure, the computer will not begin to boot up (no

lights, no sound). If there is apparently no power, be sure to check the power source and the plug at both ends: the outlet and the computer. If you are experiencing intermittent failures such as memory loss, memory corruption, or unexplained system crashes, don't rule out the power supply. It is often the culprit. Fortunately, it is easy to check and replace. Good power supplies have line-conditioning circuits, but these might not be sufficient enough in locations where the power source has substantial quality flaws. If you have problems with several systems, or if a second power supply still does not fix a related complaint, add a UPS with good line-conditioning features.

Chapter 6

Primary PC Components

OBJECTIVES

Computer Cases

Motherboards

ROM BIOS

Computer Cases

The Computer Case

To casual users, the PC is a metal container, attached by a few cables to a keyboard, mouse, and monitor. In fact, the case is more than just a box to house a computer. It often represents the identity of a specific brand of computer, and, sometimes, it is even part of the reason we purchase a particular computer. We do, after all, want something that looks good, especially if we are spending a lot of money on it.

The real value of a case does not lie in the label, color, or how pretty it looks. The case houses all the internal components, offers access to the outside world (ports and connectors), and protects the PC's delicate circuits from damage and electromagnetic interference. And it protects surrounding devices, such as TVs, from the PC's EMI.

As computer technicians, we don't usually concern ourselves with the computer case; we simply deal with whatever our customer already has. However, when it comes to recommending a computer for purchase, the size and configuration of the case should be considered. Depending on the business application, the difference between a tower and a desktop design could be important.

When considering the case, there are three general rules to follow:

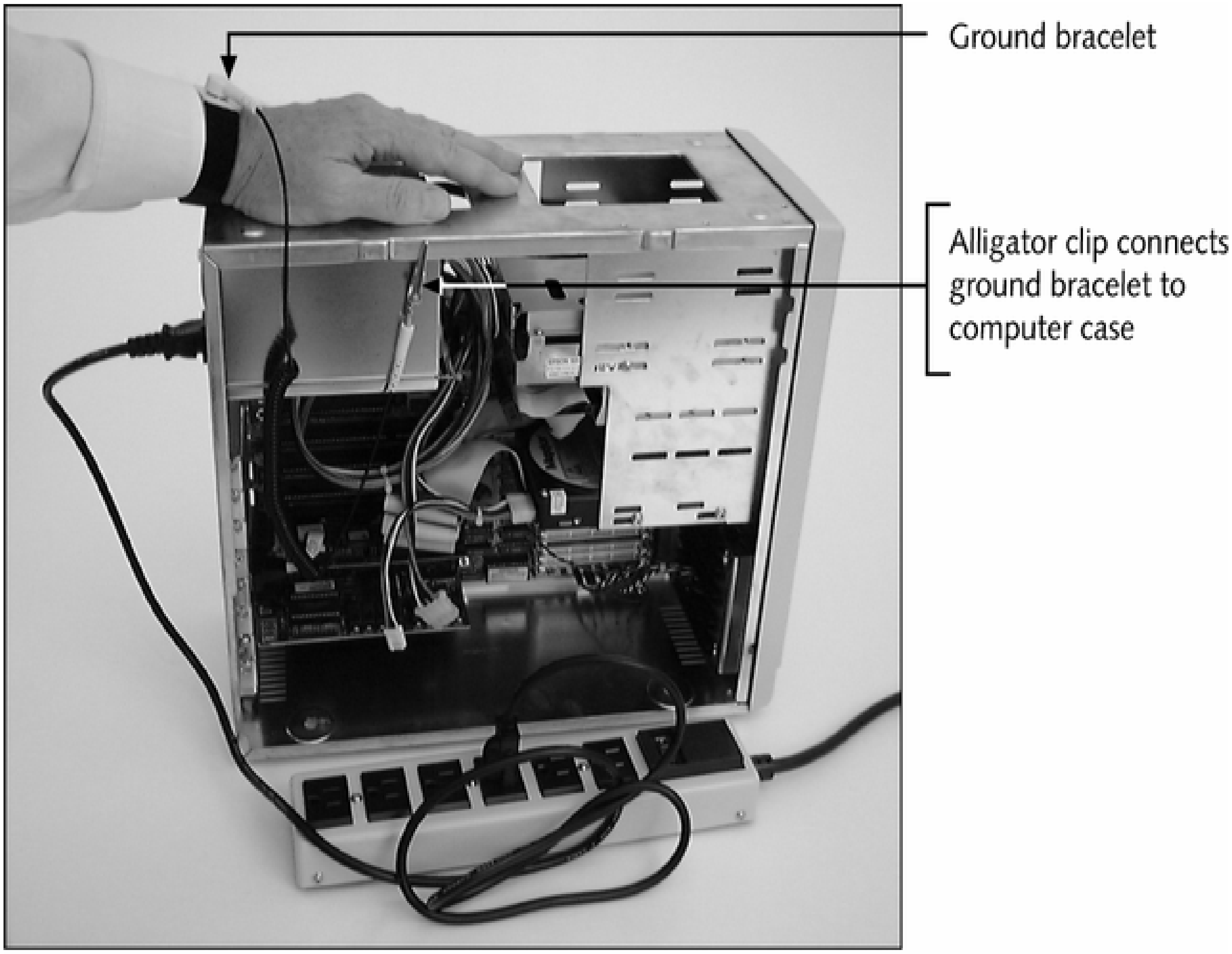
- The bigger the box, the more components it can hold (the greater the expansion potential) and, often, the better the airflow (essential for cooling). Large cases are easier to work with.
- The more compact the box, the less expansion potential it has; working on it is often much more difficult, and usually airflow is reduced.
- Smaller cases that come with a power supply usually have lower wattage, once again reducing the number of internal devices that can be installed.

Working with Cases

In any repair job that involves inspecting or replacing internal components, the technician has to open the case. That used to be very simple; the technician would remove four screws in the back of the computer with a Phillips screwdriver, then pull the case's covering shroud forward to reveal the contents. Today, however, cases come in a variety of forms, with screws in the front or back, fancy plastic bevels in front, and featuring one of several types of metal wraps—some in several parts, some in a single piece.

The majority of cases still open the old-fashioned way. But if you find yourself with one of the exceptions and can't locate screws in the back, check to see if the plastic cover in

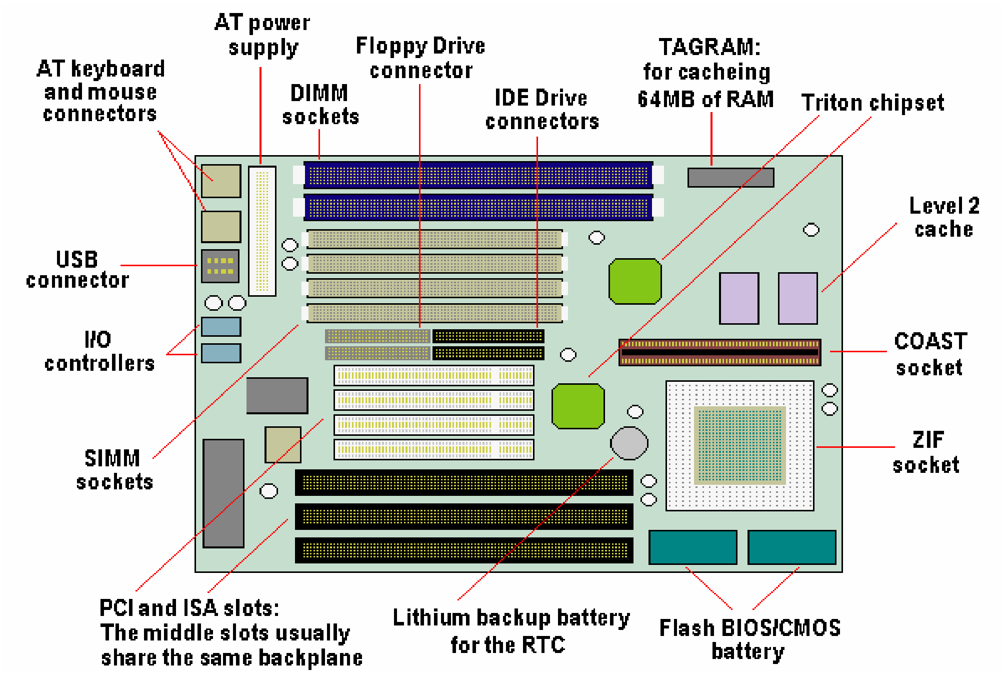
the front can be pulled off. If so, that should reveal three or four screws. Then see if the main cover can be pulled forward. If not, look for screws that secure one or more of the side panels. Some side panel designs are great for granting easy access to our next topic, motherboards. This style of case allows one to inspect or remove the motherboard without having to remove the entire outer covering.



Motherboards

The Motherboard

The motherboard defines the computer's limits of speed, memory, and expandability. A computer needs more than just a CPU and memory. To accept input from the user, it needs devices, such as a keyboard and a mouse. It also needs output devices, like monitors and sound cards, to cope with the powerful graphics and sound capabilities of the programs available today. A computer also needs "permanent" storage devices, such as floppy disk drives and hard disk drives, to store data when it is turned off. It is the function of the motherboard to provide the connectivity for all these devices, as well as for the CPU, RAM, and support ICs.



Motherboard with CPU

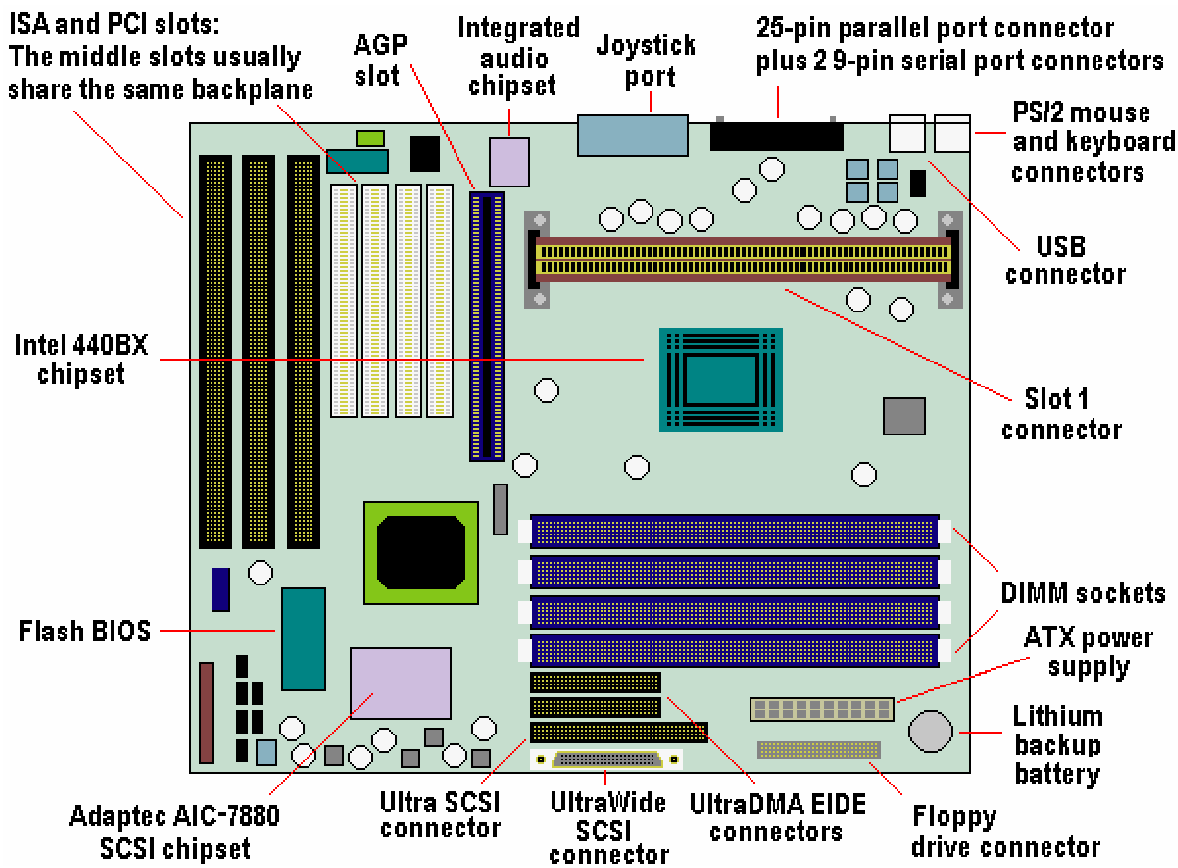
The motherboard is usually the largest circuit board found inside the computer case. Motherboards come in a variety of shapes. One size does not fit all, and careful attention to size and location of mounting holes is required before installing a new motherboard in an older computer. A motherboard needs to fit in the space allotted for it, be secure in its mounts, be properly grounded, receive sufficient ventilation (for cooling of the CPU and other heat-sensitive components), and must not conflict with other hardware. When considering the purchase of a new motherboard.

- Most "generic" motherboards will fit into "generic" computers. One good reason to consider purchasing a PC clone is that it is easier to upgrade.
- There are two major categories of motherboards: AT and ATX. The main difference between them lies in the type of power supply and main power switch each requires. When you order a new motherboard, be sure to first verify that it is compatible with the case and power supply to be used.
- If you are working on a brand-name computer, you might be required to purchase a new motherboard or other custom components from the same manufacturer.
- Before buying a motherboard, check its technical references to be sure that the new board will fit and will be compatible with any of the RAM and expansion cards the owner intends to use. Often, this information can be found in the owner's manual. If not, check the manufacturer's Web site, if one is available, or check other online resources such as technical libraries. A Web search using the keyword "motherboard" will yield sites dedicated to computer hardware.
- For all practical purposes, you cannot repair motherboards. They should be replaced if physically or electrically damaged. Your customer will get new technology, usually for a price lower than the cost of the repair.
- Because it is often the most difficult part of a system to replace (you have to remove all the equipment that is connected to it), check all other internal and external components before removing or replacing the motherboard.
- When obtaining a replacement, be sure to factor in the cost of all critical options found on the existing motherboard. Some have built-in SCSI (Small Computer System Interface) or display adapters that might not be common. In that case, either make sure the new board offers the same level of support or install the appropriate add-on card(s) to bring the system up to the existing level of operation.

Chip Sets

A motherboard comes with a variety of support chips soldered in place. The primary elements constitute the *chip set* and are designed to work with the CPU. These chips are highly complex and coordinated ICs that help the CPU manage and control the computer's system. When replacing a CPU, you must make sure that it is compatible with the chip set and supported by the motherboard. If not, the computer won't work. A basic chip set consists of a:

- Bus controller.
- Memory controller.
- Data and address buffer.
- Peripheral controller.



Motherboard with chip set

The ***form factor*** refers to the physical dimensions and size of the board and determines what type of case the board will fit into. Some of the common form factors are: -

Obsolete Form Factors

- Baby-AT, Full-size AT and LPX (semi-proprietary)

Modern Form Factors

- ATX, Micro-ATX, Flex-ATX and NLX

All Others

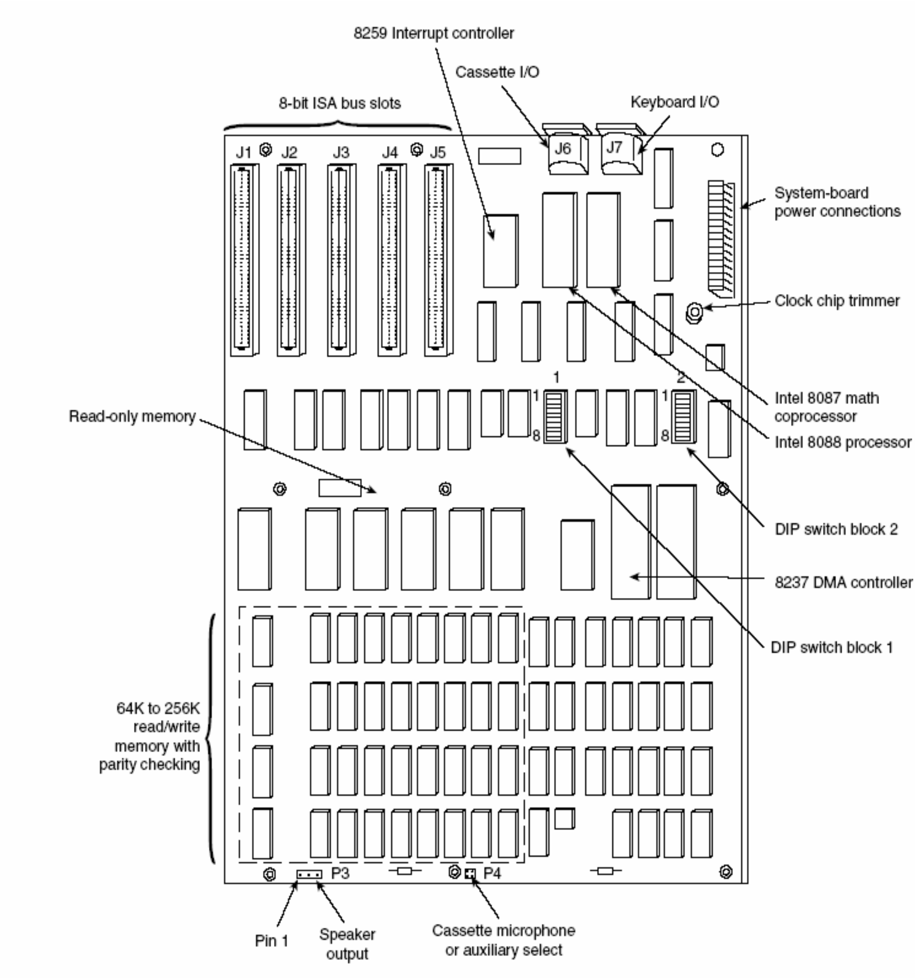
- OEM (Original Equipment Manufacturer) Designs (Compaq, HP and Acer etc)

Types of Motherboard

The different types of motherboard are discussed in details in the following sections.

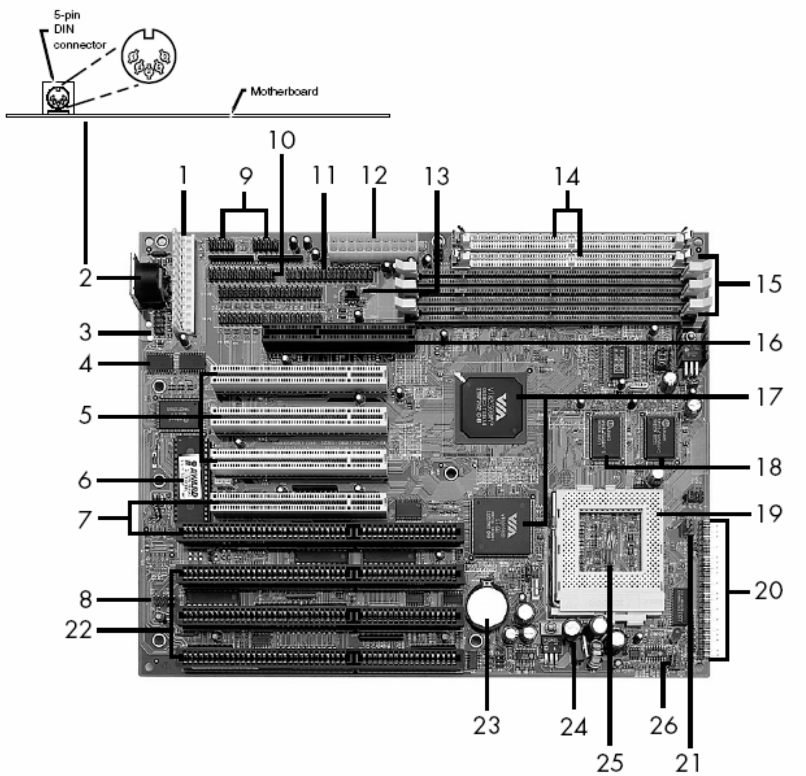
Baby-AT

The original IBM PC was released in 1981 with this motherboard. Figure 7.1 shows the board.



IBM PC motherboard

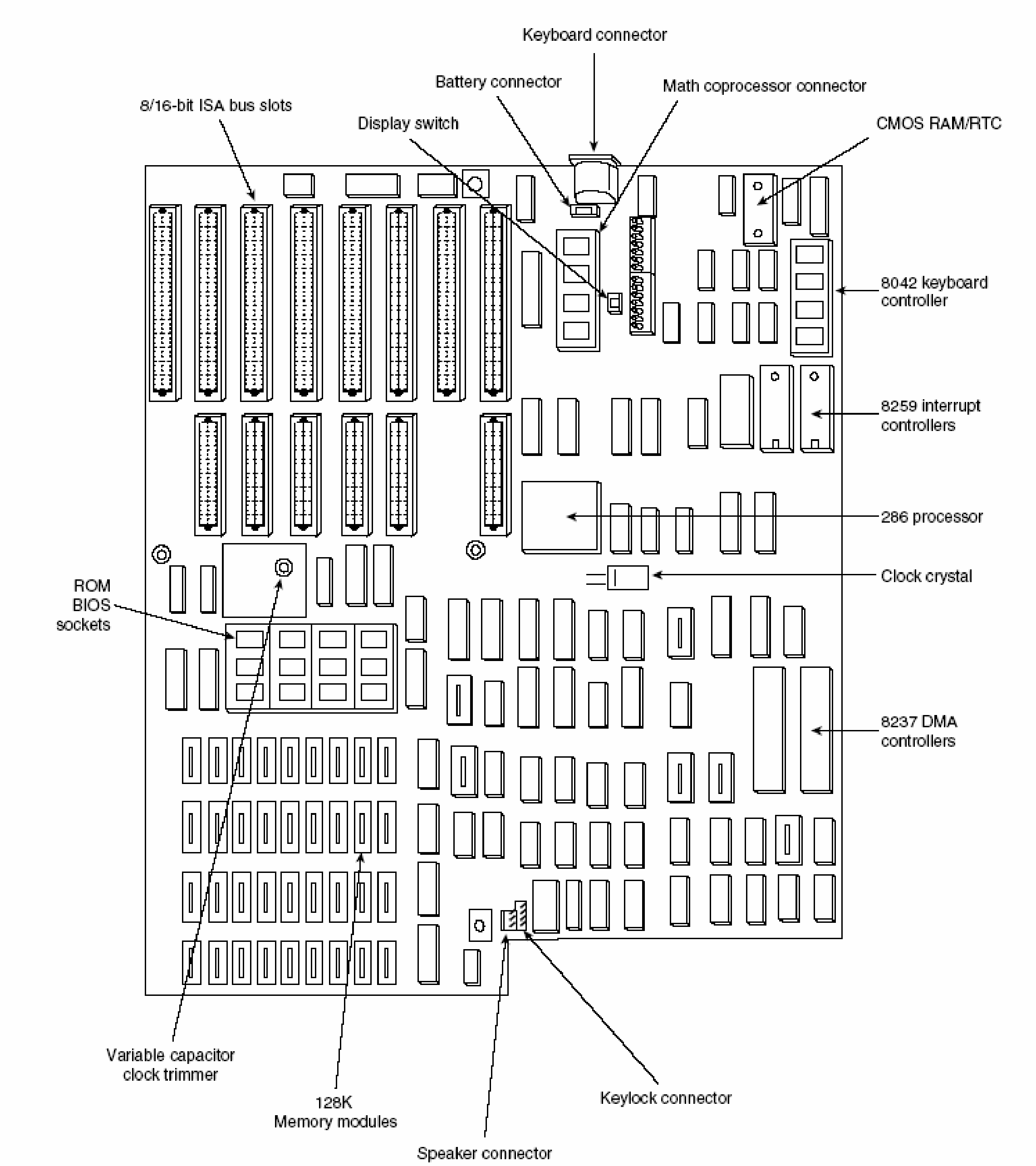
In a Baby-AT motherboard, the cards plug directly into the board at a 90° angle. It has only a keyboard connector attached directly to the board.



A Baby-AT motherboard

Full-Size AT

The full-size AT motherboard matches the original IBM AT motherboard design. This allows for a very large board of up to 12 inches wide by 13.8 inches deep. When the full-size AT board debuted in 1984, IBM needed more room for additional circuits when it migrated from the 8-bit architecture of the PC/XT to the 16-bit architecture of the AT. So, IBM started with an XT board and extended it in two directions.

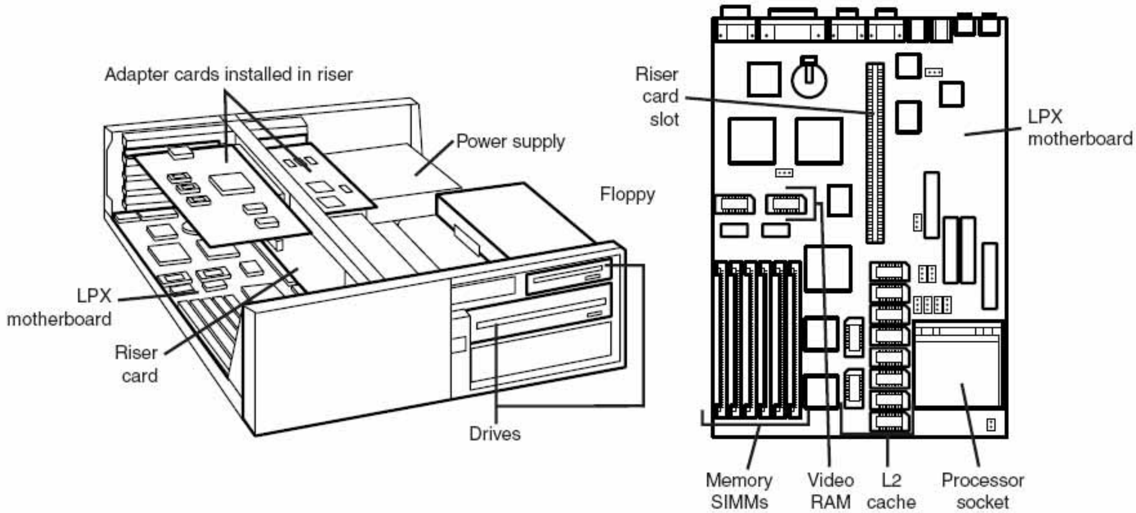


IBM AT motherboard

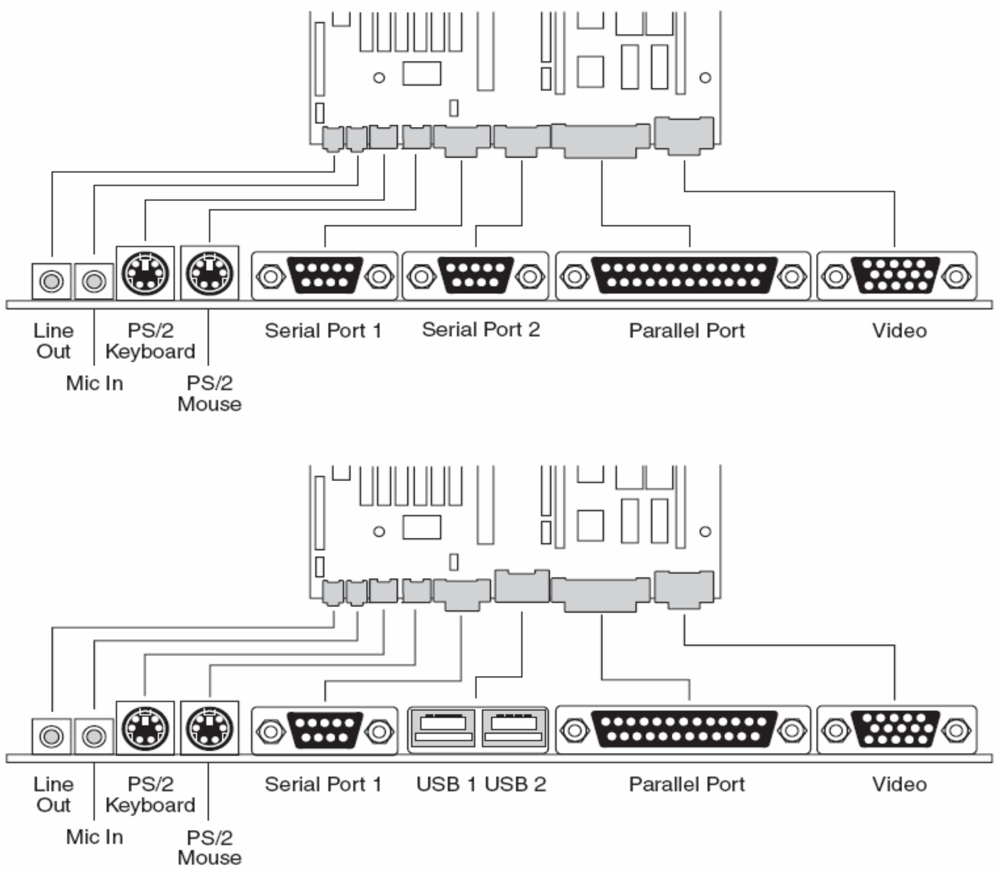
LPX

It is a semi-proprietary design developed in 1987. The LP in LPX stands for Low Profile, which is so named because these boards incorporated slots that were parallel to the main board, enabling the expansion cards to install sideways. This allowed for a slim or low-profile case design and overall a smaller system than the Baby-AT. The LPX specifications were never laid out in exact detail. Therefore these are termed as semi-proprietary and are not interchangeable between manufacturers. Some vendors, such as IBM and HP, for example, have built LPX systems that use a T-shaped riser card this

allows expansion cards to be mounted at the normal 90° angle to the motherboard but still over the motherboard.



Typical LPX system chassis and motherboard.



LPX motherboard back panel connectors.

ATX

The ATX form factor was a major progress in motherboard form factors. It has a different size compared to Baby-AT or LPX. ATX is the most popular motherboard form factor for new systems. ATX is a combination of the best features of the Baby-AT and LPX motherboard designs, with many new improvements : -

- **Built-in double high external I/O connector panel**
  The back side of the motherboard has multiple I/O connectors located directly on the board. In Baby-AT designs, there are cables running from internal connectors to the back of the case

- **Single main keyed internal power supply connector**
  Unlike Baby-AT design there is a single internal power supply connector. This connector is keyed to prevent the wrong connection, which prevents damage to the motherboard. This connector also has pins for supplying 3.3v to the motherboard, so ATX motherboards do not require built in voltage regulators that can fail easily. There are two other connectors in ATX form factor. One is Auxiliary Power connector (3.3V and 5V) and the ATX12V.

- **Relocated CPU and memory**
  The CPU and memory modules are relocated so they do not get in the way of any expansion cards and can easily be accessed for upgrade without removing any of the installed bus adapters. The CPU and memory are relocated next to the power supply, which is where the primary system fan is located.

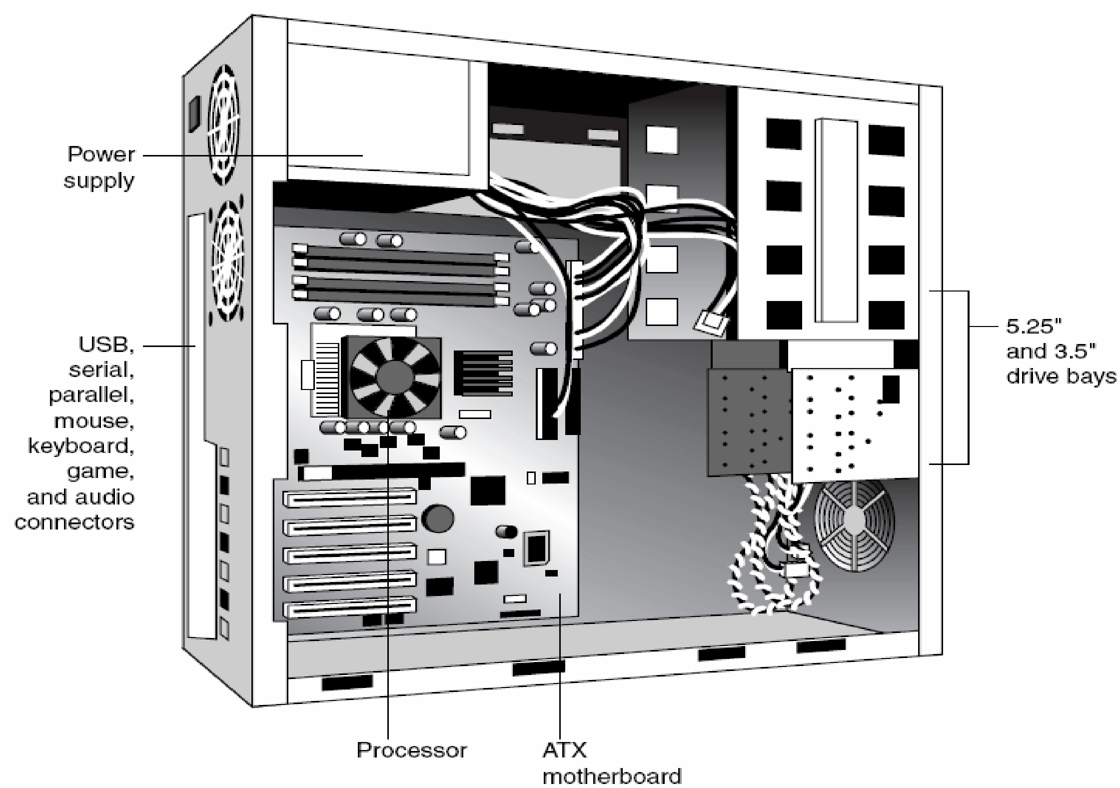
- **Relocated internal I/O connectors**
  The internal I/O connectors for the floppy and hard disk drives are relocated near the drive bays and out away from the expansion board slot and drive bay areas. This means that internal cables to the drives can be much shorter, and accessing the connectors does not require card or drive removal.

- **Improved cooling.**
  The CPU and main memory are designed and positioned to improve overall system cooling. Note that the ATX specification originally specified that the ATX power supply fan blows into the system chassis instead of outward.
  Recently, the ATX specification was revised to allow the more normal standard flow, which negatively pressurizes the case by having the fan blow outward.
  Therefore the specification allows either type of airflow; most power supply manufacturers provide ATX power supplies with fans that exhaust air from the system which is called a negative pressure design.

- **Lower cost to manufacture**
  Following are some simplifications which help to reduce the cost of making an ATX motherboard: -

  - No of cables to external port connectors found on Baby-AT motherboards
  - No onboard 3.3v voltage regulators.
  - Shorter internal drive cables
  - No cables for standard external serial or parallel ports.

The picture below shows the new ATX system layout. Notice how the processor is positioned near the power supply.

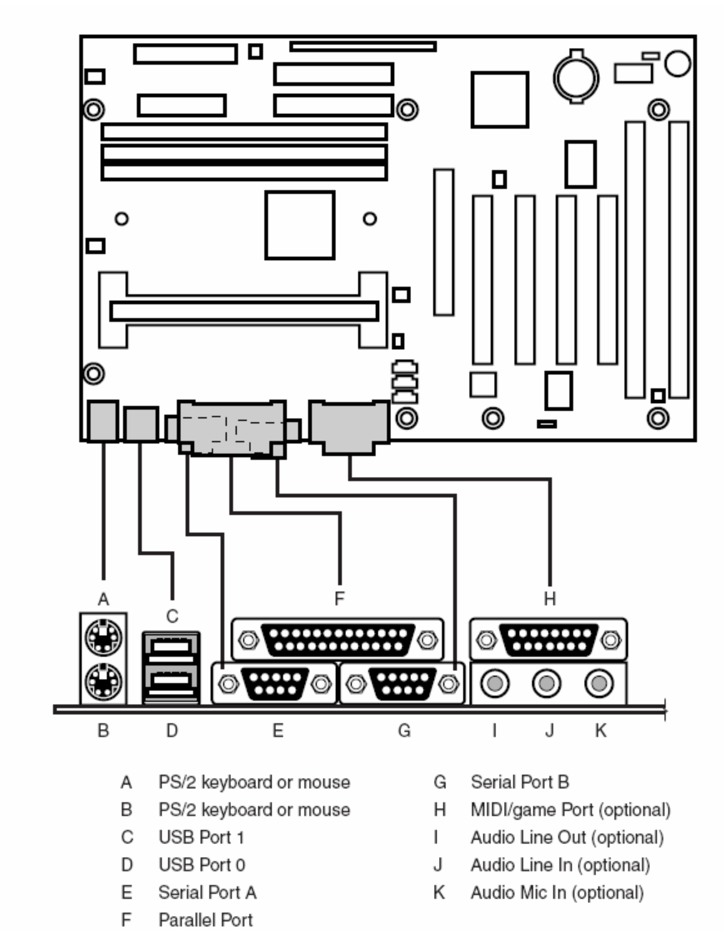


The ATX motherboard is oriented so that the CPU socket is near the power supply fan and case fan.

There are actually two basic sizes of standard ATX boards. In addition to a full-size ATX layout, Intel also has specified a mini-ATX design, which is a fully compatible subset in size and will fit into the same case:

- A full-size ATX board is 12” wide × 9.6” deep (305mm×244mm).
- The mini-ATX board is 11.2” × 8.2” (284mm×208mm).

In addition, two smaller variations of ATX exist, called Micro-ATX and Flex-ATX. Clearly, the advantages of the ATX form factor make it the best choice for new systems. The best way to tell whether your system has an ATX-board design without removing the lid is to look at the back of the system. Two distinguishing features identify ATX. One is that the expansion boards plug directly into the motherboard. There is usually no riser card as with LPX or NLX, so the slots are perpendicular to the plane of the motherboard. Also, ATX boards have a unique double-high connector area for all the built-in connectors on the motherboard.



Typical ATX motherboard rear panel connectors.

Micro-ATX

It was introduced by Intel in December of 1997 as an upgrade of the ATX form factor for smaller and lower-cost systems. The reduced size as compared to standard ATX allows for a smaller chassis, motherboard, and power supply, reducing the cost of entire system. The micro-ATX form factor is also backward compatible with the ATX form factor and can be used in full-size ATX cases. Of course, a micro-ATX case cannot fit a full-size

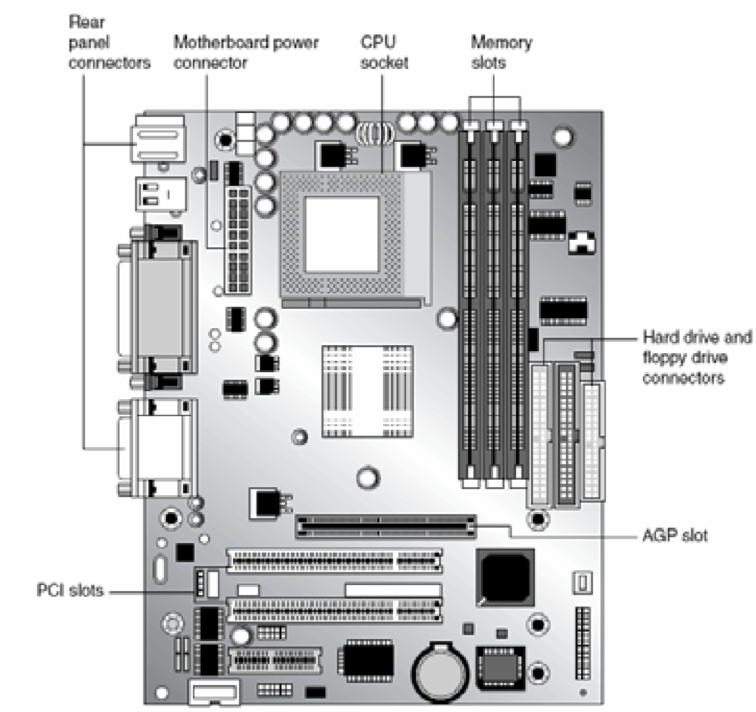
ATX board. The main differences between micro-ATX and standard or mini-ATX are as follows:

- Reduced width motherboard (9.6” instead of 12” or 11.2”)
- Fewer I/O bus expansion slots (four maximum, although most boards feature only three)
- Smaller power supply optional(SFX form factor)

The micro-ATX form factor is similar to ATX for compatibility. The similarities include the following:

- Standard ATX 20-pin power connector
- Standard ATX I/O panel
- Mounting holes are a subset of ATX

Because of above similarities a micro-ATX motherboard will easily work in a standard ATX chassis with a standard ATX power supply, as well as the smaller micro-ATX chassis and SFX power supply. A typical micro-ATX motherboard is shown in the picture below.



A typical micro-ATX motherboard’s dimensions are 9.6”×9.6”.

Flex-ATX

Intel added the flex-ATX which is smaller than the micro-ATX specification in 1999. This smaller design is planned to allow less expensive and smaller systems.

Flex-ATX board is 9.0” × 7.5” and is the smallest of the ATX boards. It has sockets like: -

- Socket 7

- Socket A for AMD processors
- Socket 370 for the PPGA (Plastic Pin Grid Array)
- FCPGA (Flip Chip PGA)

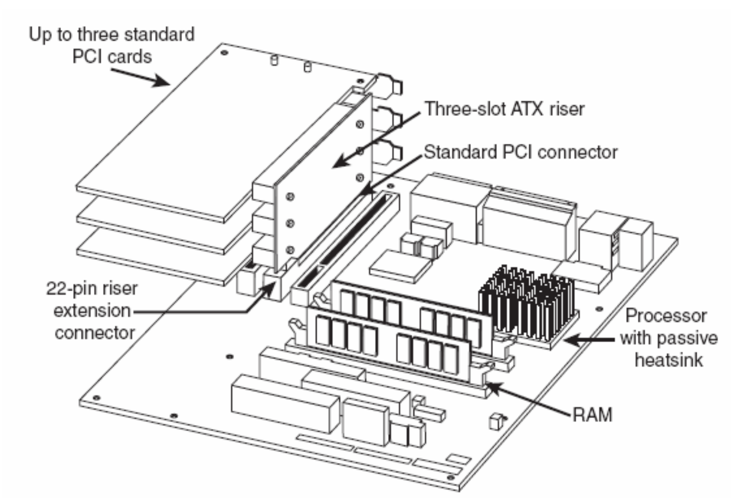
Because both Intel and AMD use socketed processors for their latest designs, so there is no compatibility problem.

Maximum dimensions of ATX Motherboard Form Factors

Form Factor Max.	**Width Max.**	**Depth**
ATX	12.0”	9.6”
Mini-ATX	11.2”	8.2”
Micro-ATX	9.6”	9.6”
Flex-ATX	9.0”	7.5”

ATX Riser

In December 1999, Intel introduced a riser card design modification for ATX motherboards. The design includes the addition of a 22-pin (2×11) connector to one of the PCI slots on the motherboard, along with a two- or three-slot riser card that plugs in. The riser enables two or three PCI cards to be installed. The riser does not support AGP.

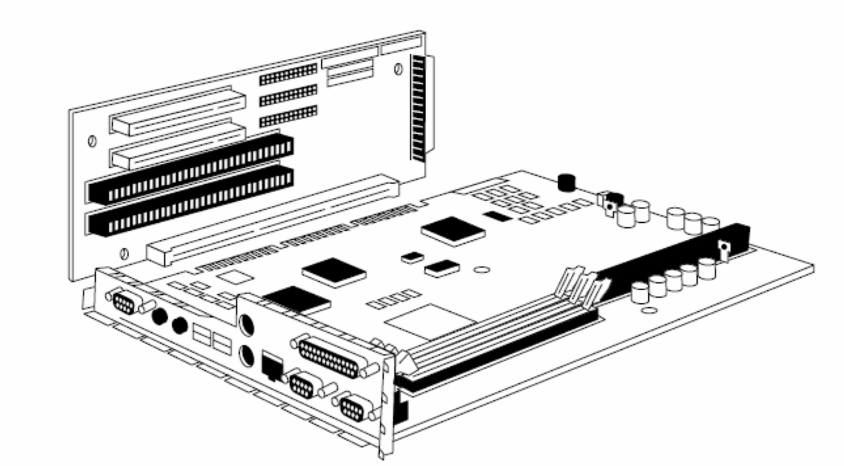


Three-slot ATX riser implementation on a micro-ATX motherboard.

NLX

NLX is a low-profile form factor designed to replace the nonstandard LPX design used in previous low-profile systems. It was introduced by Intel 1996 by Intel. It has become the form factor of choice for slim line corporate desktop systems. NLX is similar in initial appearance to LPX, but with numerous improvements designed to enable full integration of the latest technologies. NLX is basically an improved version of the proprietary LPX

design, but, unlike LPX, NLX is fully standardized, which means you should be able to replace one NLX board with another from a different manufacturer—something that was not possible with LPX.



NLX motherboard and riser combination

Proprietary Designs

Motherboards that are **not** one of the standard form factors, such as full-size or Baby-AT, ATX, mini-ATX, micro-ATX, flex-ATX, or NLX, are considered **proprietary**. LPX systems fall into this category. Most people purchasing PCs should avoid proprietary designs because they do not allow for a future other board, power supply, or case upgrade, which limits future use and serviceability of the system.

Motherboard Components

A modern motherboard has several components built in, including various sockets, slots, connectors, chips, and so on. This section examines the components found on a typical motherboard. Most modern motherboards have at least the following major components on them:

- Processor socket/slot
- Chipset
- Super I/O chip
- ROM BIOS
- RAM Memory Sockets
- Motherboard Bus Slots
- CPU voltage regulator
- Battery

Chipsets

The chipset is the main component of a motherboard. Two boards with the same chipsets are functionally same. The chipset contains the processor bus interface which is called front-side bus or FSB, memory controllers, bus controllers, I/O controllers, and more.
The chipset has the motherboard circuits. The chipset connects the processor with everything else. The processor cannot talk to the memory, adapter boards, devices, and so on without going through the chipset. The chipset is the main hub and central nervous system of the PC. If you think of the processor as the brain, the chipset is the spine and central nervous system.
Chipset determines: -

- What type of processor type of processor you can fit on the motherboard?
- How fast the buses will run?
- What type, speed and amount of memory you can use?

Chipset can be more important even than the processor. Systems with slower processor but a better chipset can perform better than a system with faster processors. It is just like how a car with less power might win a race through better cornering and braking. It is recommended to start with choosing the chipset first because the chipset decision then determines the processor, memory, I/O, and expansion capabilities.

Super I/O Chips

The third major chip seen on many PC motherboards is called the Super I/O chip. This chip normally integrates devices formerly found on separate expansion cards in older systems.

Most Super I/O chips contain, at a minimum, the following components:
- Floppy controller
- Dual serial port controllers
- Parallel port controller

On the newer motherboards the role of this chip has decreased and more functions are now integrated into the main chipsets. There are basically two reasons for that: -
- You get better motherboard bus speed by above said function integration
- There is an increased tendency of emerging USB devices, reducing the dependence on the serial and parallel interface devices

Mother Board Buses

A bus is a common pathway across which data can travel within a computer. This pathway is used for communication and can be established between two or more

computer elements. The PC has a hierarchy of different buses. Most modern PCs have at least three buses; some have four or more. They are hierarchical because each slower bus is connected to the faster one above it. Each device in the system is connected to one of the buses, and some devices (primarily the chipset) act as bridges between the various buses.

The Processor Bus

The processor bus (also called the Front-side bus or FSB) is the communication pathway between the CPU and Motherboard chipset (Memory Controller). This bus runs at the full motherboard speed—normally between 66MHz and 400MHz in modern systems, depending on the particular board and chipset design. This same bus also transfers data between the CPU and an external (L2) memory cache on Socket-7 (Pentium class) systems.

The Memory Bus

The memory bus is used to transfer information between the CPU and main memory—the RAM in your system. This bus is connected to the Memory Controller. Depending on the type of memory your chipset (and therefore motherboard) is designed to handle, the memory controller runs the memory bus at various speeds. The best solution is if the memory bus runs at the same speed as the processor bus. Systems that use PC133 SDRAM have a memory bandwidth of 1,066MB/sec, which is the same as the 133MHz CPU bus. In another example, Athlon and Pentium III systems running a 266MHz processor bus also run PC2100 DDR-SDRAM, which has a bandwidth of 2,133MB/sec—exactly the same as the processor bus in those systems. Also, systems running a Pentium 4 with its 400MHz processor bus also use dual-channel RDRAM memory, which runs 1,600MB/sec for each channel, or a combined bandwidth (both memory channels run simultaneously) of 3,200MB/sec, which is exactly the same as the Pentium 4 CPU bus.

Running memory at the same speed as the processor bus negates the need for having cache memory on the motherboard. That is why when the L2 cache moved into the processor, nobody added an L3 cache to the motherboard.

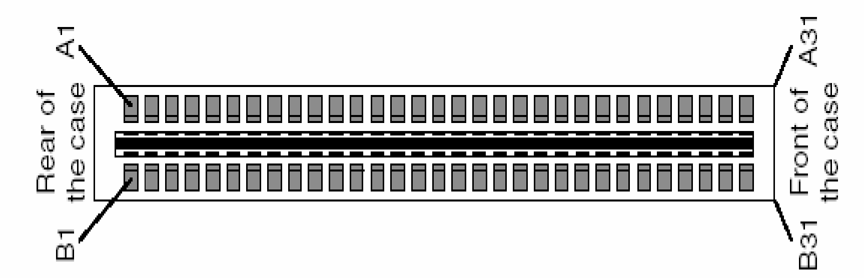
The ISA Bus

ISA, which is an abbreviation for Industry Standard Architecture, is the bus architecture that was introduced as an 8-bit bus with the original IBM PC in 1981; it was later expanded to 16 bits with the IBM PC/AT in 1984. ISA is the basis of the modern personal computer and the primary architecture used in the vast majority of PC systems on the market today. It might seem amazing that such a presumably out of date architecture is used in today’s high-performance systems, but this is true for reasons of

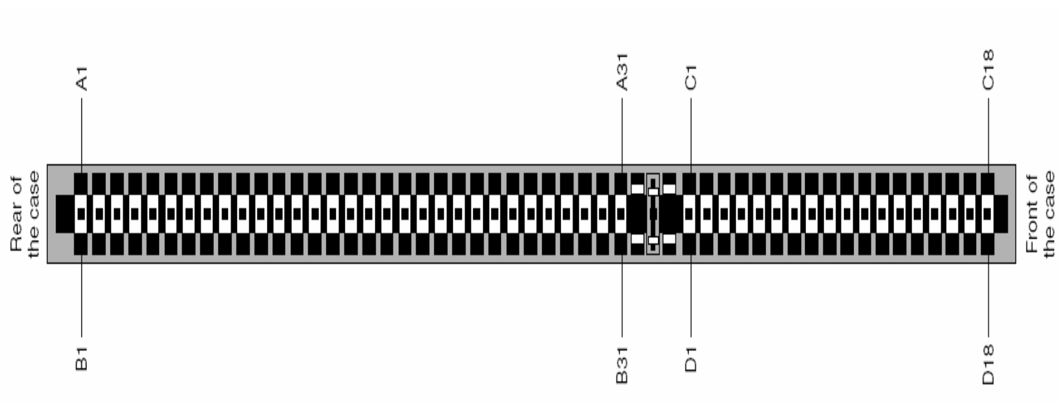
reliability, affordability, and compatibility, plus this old bus is still faster than many of the peripherals we connect to it.

Two versions of the ISA bus exist, based on the number of data bits that can be transferred on the bus at a time. The older version is an 8-bit bus; the newer version is a 16-bit bus. The original 8-bit version ran at 4.77MHz in the PC and XT. The 16-bit version used in the AT ran at 6MHz and then 8MHz. Later, the industry as a whole agreed on an 8.33MHz maximum standard speed for 8/16-bit versions of the ISA bus for backward compatibility. Some systems have the capability to run the ISA bus faster than this, but some adapter cards will not function properly at higher speeds. ISA data transfers require anywhere from two to eight cycles. Therefore, the theoretical maximum data rate of the ISA bus is about 8MB/sec, as the following formula shows:

The bandwidth of the 8-bit bus would be half this figure (4.17MB/sec). Remember, however, that these figures are theoretical maximums; because of I/O bus protocols, the effective bandwidth is much lower—typically by almost half. Even so, at about 8MB/sec, the ISA bus is still faster than many of the peripherals connected to it, such as serial ports, parallel ports, floppy controllers, keyboard controllers, and so on.



The 8-bit ISA bus connector.



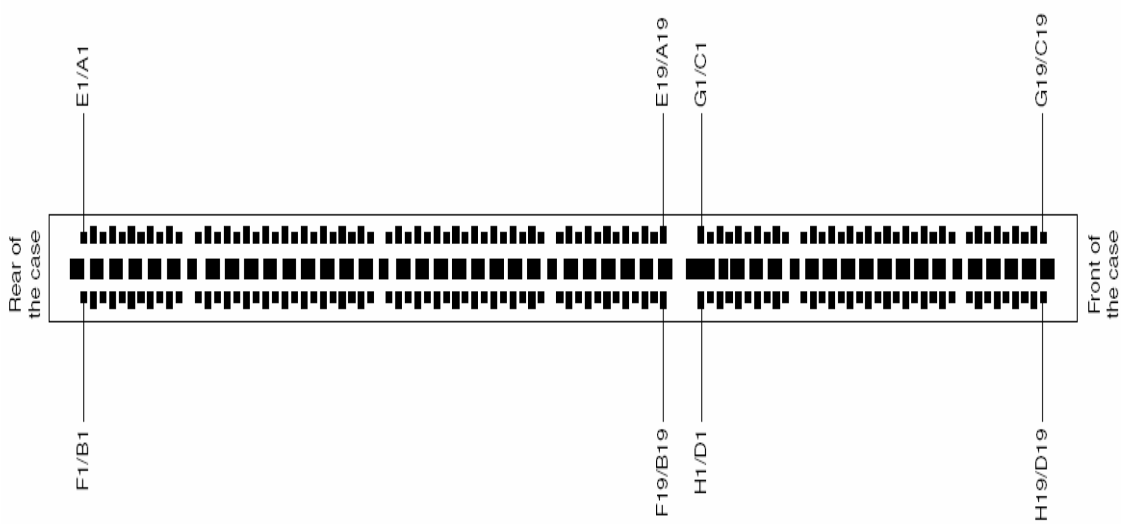
The ISA 16-bit bus connector.

The EISA Bus

EISA is an abbreviation for Extended Industry Standard Architecture. This standard was announced in September 1988 as a response to IBM's introduction of the MCA bus—more specifically, to the way that IBM wanted to handle licensing of the MCA bus. Vendors did not feel obligated to pay retroactive royalties on the ISA bus, so they turned their backs on IBM and created their own buses.

The EISA standard was developed primarily by Compaq and was intended to be its way of taking over future development of the PC bus from IBM. Compaq knew that nobody would clone its bus if it was the only company that had it, so it essentially gave the design away to other leading manufacturers. Compaq formed the EISA committee, a nonprofit organization designed specifically to control development of the EISA bus. Very few EISA adapters were ever developed. Those that were developed centered mainly around disk array controllers and server-type network cards. The EISA bus was essentially a 32-bit version of ISA. Unlike the MCA bus from IBM, you could still use older 8-bit or 16-bit ISA cards in 32-bit EISA slots, providing for full backward compatibility. As with MCA, EISA also allowed for automatic configuration of EISA cards via software.

The EISA bus added 90 new connections (55 new signals plus grounds) without increasing the physical connector size of the 16-bit ISA bus. At first glance, the 32-bit EISA slot looks a lot like the 16-bit ISA slot. The EISA adapter, however, has two rows of stacked contacts. The first row is the same type used in 16-bit ISA cards; the other, thinner row extends from the 16-bit connectors. Therefore, ISA cards can still be used in EISA bus slots.



The EISA bus connector.

Although this compatibility was not enough to ensure the popularity of EISA buses, it is a feature that was carried over into the VL-Bus standard that followed.

The EISA bus can handle up to 32 bits of data at an 8.33MHz cycle rate. Most data transfers require a minimum of two cycles, although faster cycle rates are possible if an adapter card provides tight timing specifications.

Local Buses

The I/O buses discussed so far (ISA, MCA, and EISA) have one thing in common: relatively slow speed. The next three bus types that are discussed in the following few sections all use the *local bus* concept explained in this section to address the speed issue. The three main local buses found in PC systems are

- VL-Bus (VESA local bus)
- PCI
- AGP

The speed limitation of ISA, MCA, and EISA is a carryover from the days of the original PC, when the I/O bus operated at the same speed as the processor bus. As the speed of the processor bus increased, the I/O bus realized only nominal speed improvements, primarily from an increase in the bandwidth of the bus. The I/O bus had to remain at a slower speed because the huge installed base of adapter cards could operate only at slower speeds.

VESA Local Bus

The Video Electronics Standards Association (VESA) local bus was the most popular local bus design from its debut in August 1992 through 1994. It was created by the VESA committee, a nonprofit organization originally founded by NEC to further develop video display and bus standards. In a similar fashion to how EISA evolved, NEC had done most of the work on the VL-Bus (as it would be called) and, after founding the nonprofit VESA committee, NEC turned over future development to VESA. At first, the local bus slot seemed primarily designed to be used for video cards. Improving video performance was a top priority at NEC to help sell its high-end displays as well as its own PC systems. By 1991, video performance had become a real bottleneck in most PC systems.

The VL-Bus can move data 32 bits at a time, enabling data to flow between the CPU and a compatible video subsystem or hard drive at the full 32-bit data width of the 486 chip. The maximum rated throughput of the VL-Bus is 133MB/sec. In other words, local bus went a long way toward removing the major bottlenecks that existed in earlier bus configurations.

The PCI Bus

In early 1992, Intel spearheaded the creation of another industry group. It was formed with the same goals as the VESA group in relation to the PC bus. Recognizing the need to overcome weaknesses in the ISA and EISA buses, the PCI Special Interest Group was formed. The PCI bus specification was released in June 1992 as version 1.0 and since then has undergone several upgrades. PCI redesigned the traditional PC bus by inserting another bus between the CPU and the native I/O bus by means of bridges. Rather than tap directly into the processor bus, with its delicate electrical timing (as was done in the VL-Bus), a new set of controller chips was developed to extend the bus.

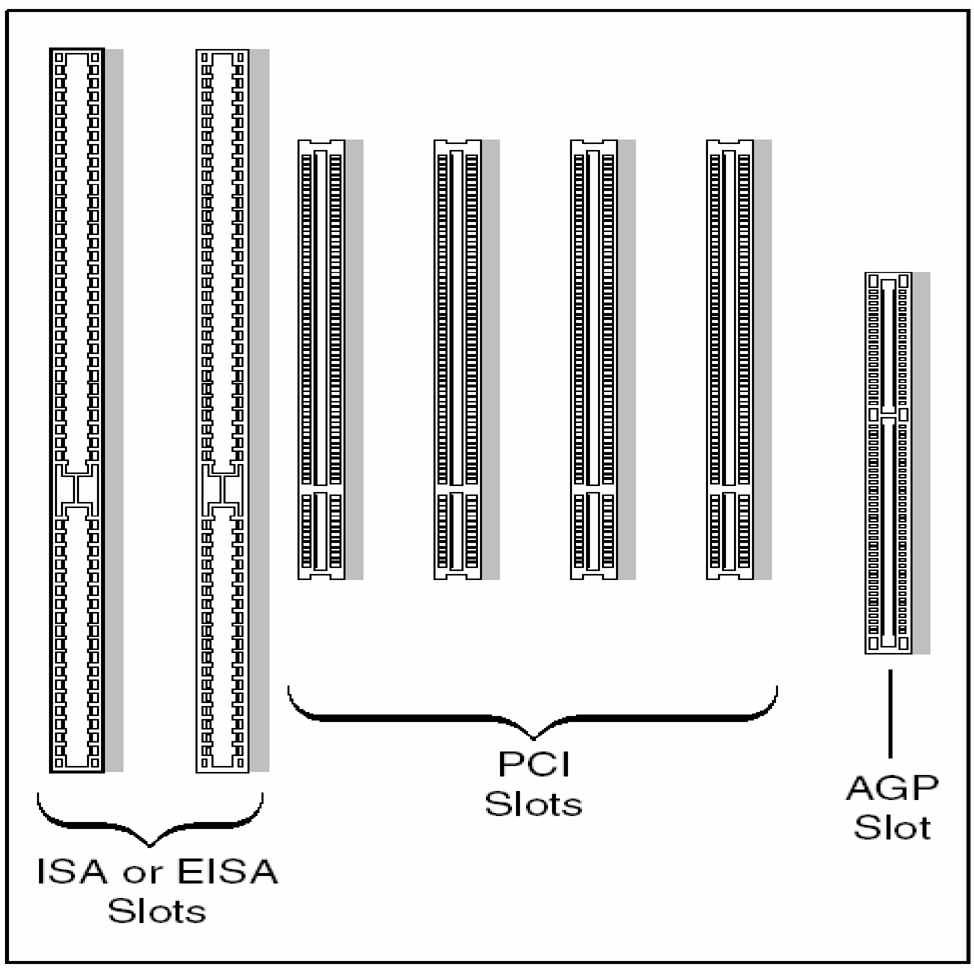
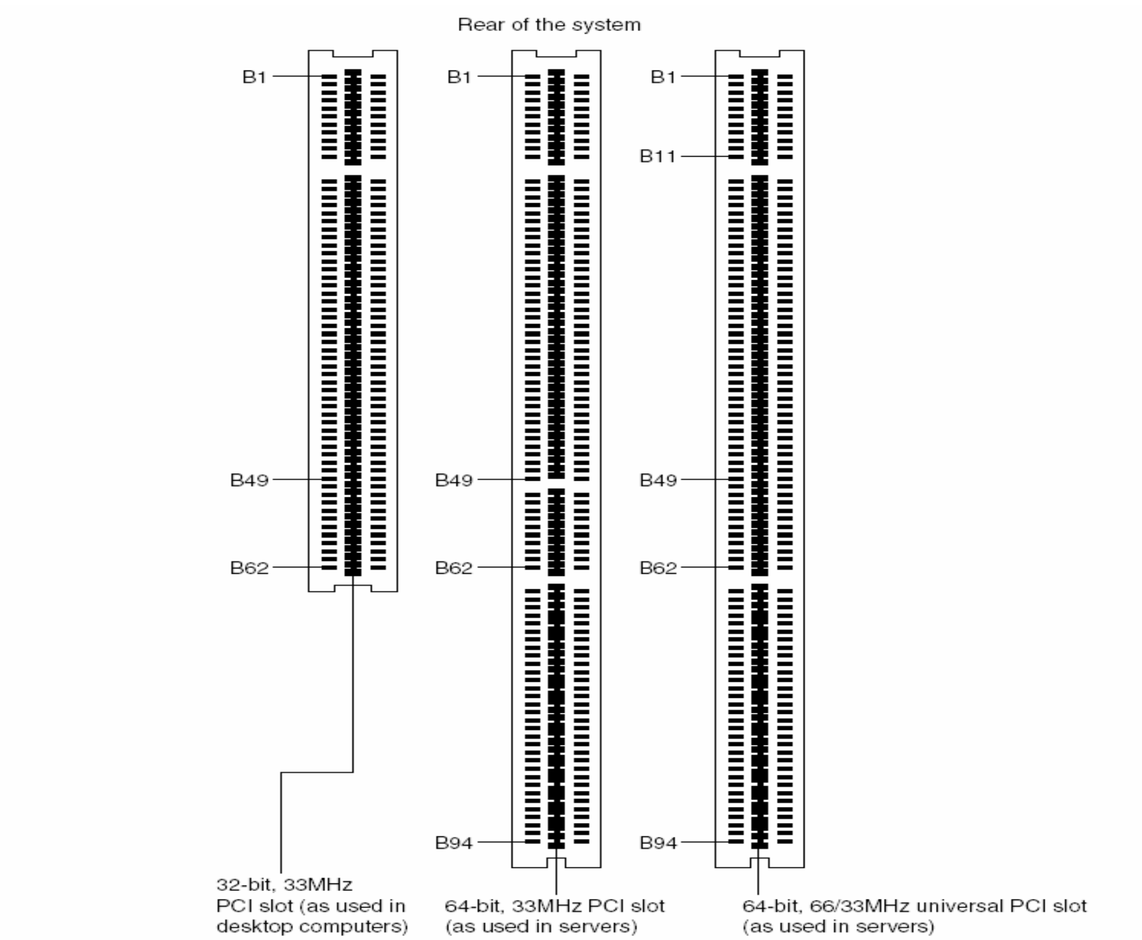
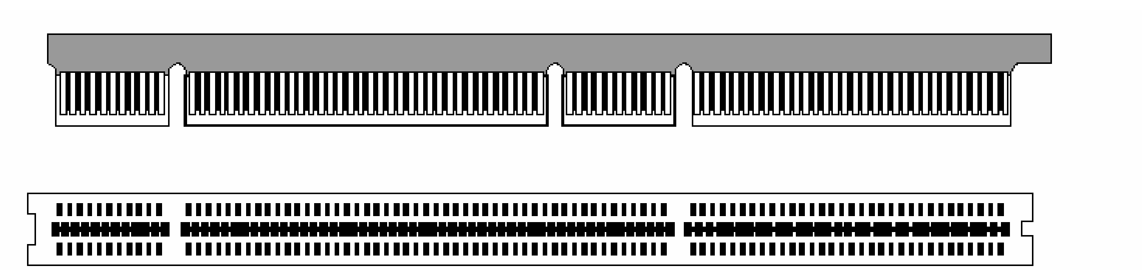


Figure 6. 1 Typical configuration of 32-bit 33MHz PCI slots in relation to ISA or EISA and AGP slots.



A 32-bit, 33MHz PCI slot (left) compared to a 64-bit 33MHz PCI slot (center) and a 64-bit universal PCI slot that runs at 66MHz (right).



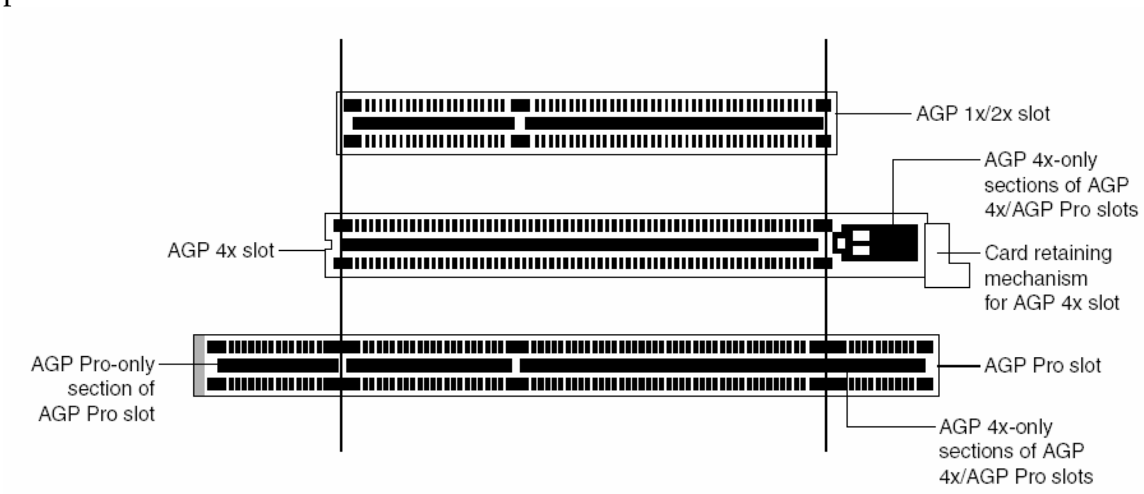
A 64-bit universal PCI card (top) compared to the 64-bit universal PCI slot (bottom)

Accelerated Graphics Port

The AGP was created by Intel as a new bus specifically designed for high-performance graphics and video support. AGP is based on PCI, but it contains a number of additions and enhancements and is physically, electrically, and logically independent of PCI. For example, the AGP connector is similar to PCI, although it has additional signals and is positioned differently in the system. Unlike PCI, which is a true bus with multiple

connectors (slots), AGP is more of a point-to-point high-performance connection designed specifically for a video card in a system because only one AGP slot is allowed for a single video card. The AGP specification 1.0 originally was released by Intel in July of 1996 and defined a 66MHz clock rate with 1x or 2x signaling using 3.3 volts. AGP version 2.0 was released in May 1998 and added 4x signaling as well as a lower 1.5v operating capability.

The standard AGP 1x/2x, AGP 4x, and AGP Pro slots are compared to each other in the pictures. The latest revision to the AGP specification for PCs is AGP 8x, which was announced in August 2000; it defines a faster 8x transfer mode for even greater performance than before.



AGP standard (1x/2x), AGP 4x, and AGP Pro slots compared to each other. AGP 4x and AGP Pro can accept AGP 1x, 2x, and 4x cards.

AGP is a high-speed connection and runs at a base frequency of 66MHz (actually 66.66MHz), which is double that of standard PCI. In the basic AGP mode, called 1x, a single transfer is done every cycle. Because the AGP bus is 32 bits (4 bytes) wide, at 66 million times per second it would be capable of transferring data at a rate of about 266MB/sec! The original AGP specification also defines a 2x mode, in which two transfers are performed every cycle, resulting in 533MB/sec.

System Resources

System resources are the communications channels, addresses, and other signals used by hardware devices to communicate on the bus. At their lowest level, these resources typically include the following:
Memory addresses

- IRQ (interrupt request) channels
- DMA (direct memory access) channels

- I/O port addresses

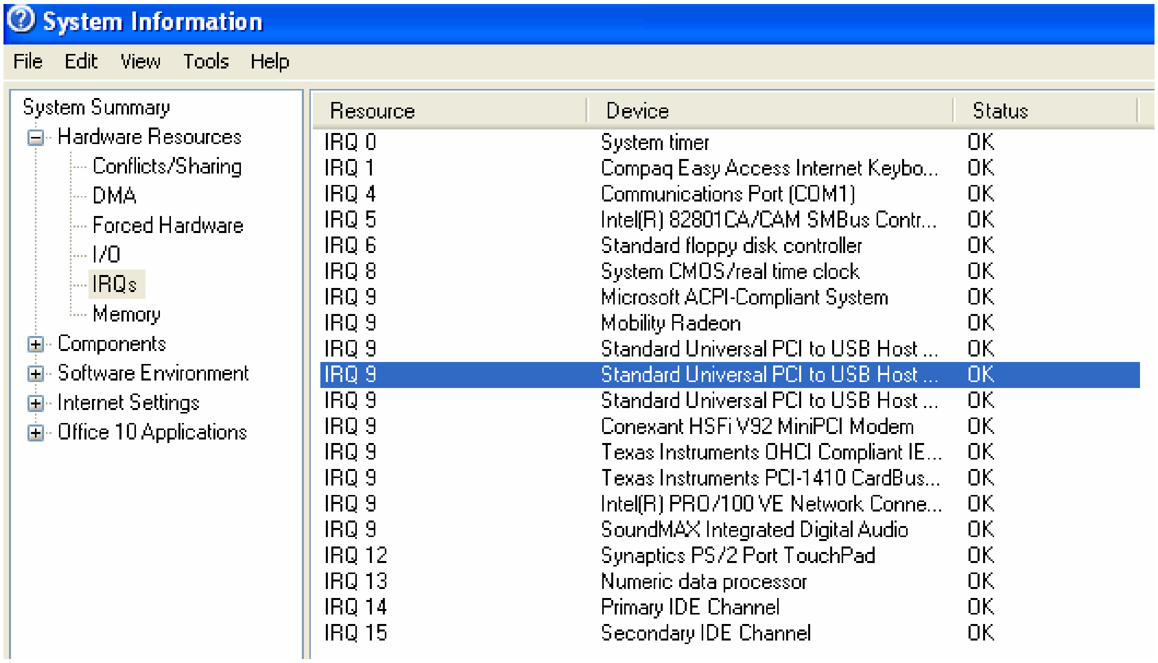
Memory conflicts are perhaps the most troublesome of these and certainly the most difficult to fully explain and overcome. IRQs cause more problems than DMA because they are in much higher demand; virtually all cards use IRQ channels. Fewer problems exist with DMA channels because fewer cards use them, and there are usually more than enough channels to go around. I/O ports are used by all hardware devices on the bus, but there are technically 64KB of them, which means there are plenty to go around. With all these resources, you have to ensure that a unique card or hardware function uses each resource; in most cases they cannot or should not be shared.

These resources are required and used by many components of your system. Adapter cards need these resources to communicate with your system and accomplish their purposes. Not all adapter cards have the same resource requirements. A serial communications port, for example, needs an IRQ channel and I/O port address, whereas a sound card needs these resources and normally at least one DMA channel as well. Most network cards use an IRQ channel and an I/O port address, and some might also use a 16KB block of memory addresses.

Interrupts

These are the hardware lines over which devices can send signals to get the attention of the processor when the device is ready to accept or send information. Each device must have a unique IRQ line. An interrupt is a request for attention from the processor. When the processor receives an interrupt, it suspends its current operations, saves the status of its work, and transfers control to a special routine known as an interrupt handler, which contains the instructions for dealing with the particular situation that caused the interrupt.

Hardware interrupts are generally prioritized by their numbers; with some exceptions, the highest priority interrupts have the lowest numbers. Higher-priority interrupts take precedence over lower-priority interrupts by interrupting them. As a result, several interrupts can occur in your system concurrently, each interrupt nesting within another. Note that the IRQ for devices using ISA buses cannot be shared while for the devices using PCI bus can share the same IRQ. To find out the IRQ values click Start, Programs, Accessories, System Tools and System information. Then expand the Hardware resources so that you can see the IRQ information.



System IRQs for a typical PC

IRQ Conflicts

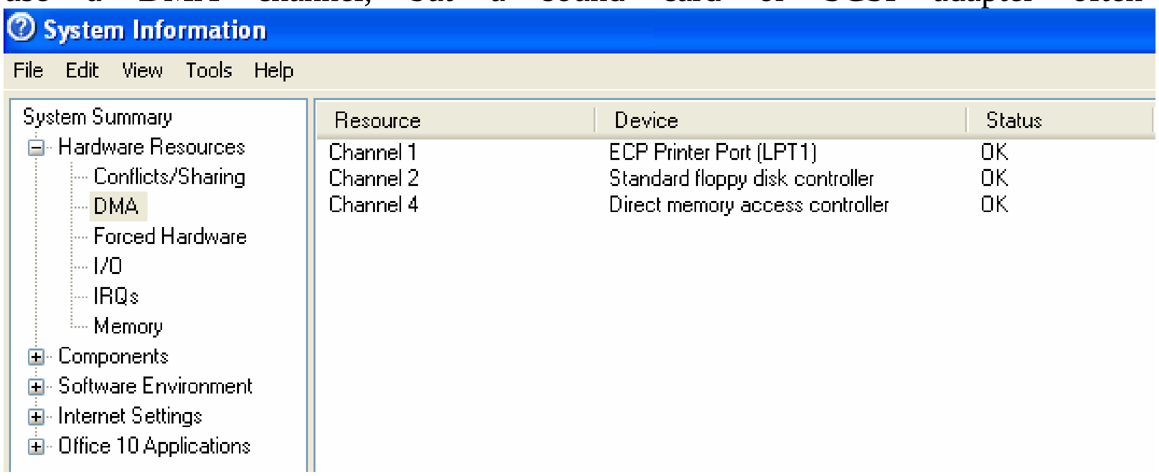
Generally there is an IRQ conflict between the integrated COM2 and an internal (card-based) modem. The base for this problem is that that true PC card–based modems (not the so-called Win Modems, which are software based) use a serial port as part of their circuit. This serial port is set as COM2: by default. Your PC sees this as having two COM2: ports, each using the same IRQ and I/O port address resources. To solve this problem, enter the system BIOS Setup and disable the built-in COM2 port in the system. Another common IRQ conflict involves serial (COM) ports. You might have noticed in the preceding two sections that two IRQs are set aside for two COM ports. IRQ 3 is used for COM2:, and IRQ 4 is used for COM1. The problem occurs when you have more than two serial ports in a system. When people add COM3: and COM4: ports, they often don't set them to nonconflicting interrupts, which results in a conflict and the ports not working.

If a device listed in the table is not present, such as the motherboard mouse port (IRQ 12) or parallel port 2 (IRQ 5), you can consider those interrupts as available. For example, a second parallel port is a rarity, and most systems have a sound card installed and set for IRQ 5. Also, on most systems IRQ 15 is assigned to a secondary IDE controller. If you do not have a second IDE hard drive, you could disable the secondary IDE controller to free up that IRQ for another device. Note that an easy way to check your interrupt settings is to use the Device Manager in Windows
95/98, NT or Windows 2000. By double-clicking the Computer Properties icon in the Device Manager, you can get concise lists of all used system resources. Microsoft has

also included a program called HWDIAG on Windows 95B; Windows 98 and above feature the System Information program.

DMA Channels

Memory access that does not involve the microprocessor. DMA is frequently used for data transfer directly between memory and a peripheral device such as a disk drive.
Direct Memory Access (DMA) channels are used by high-speed communications devices that must send and receive information at high speed. A serial or parallel port does not use a DMA channel, but a sound card or SCSI adapter often does.



System DMA Channel for a typical PC

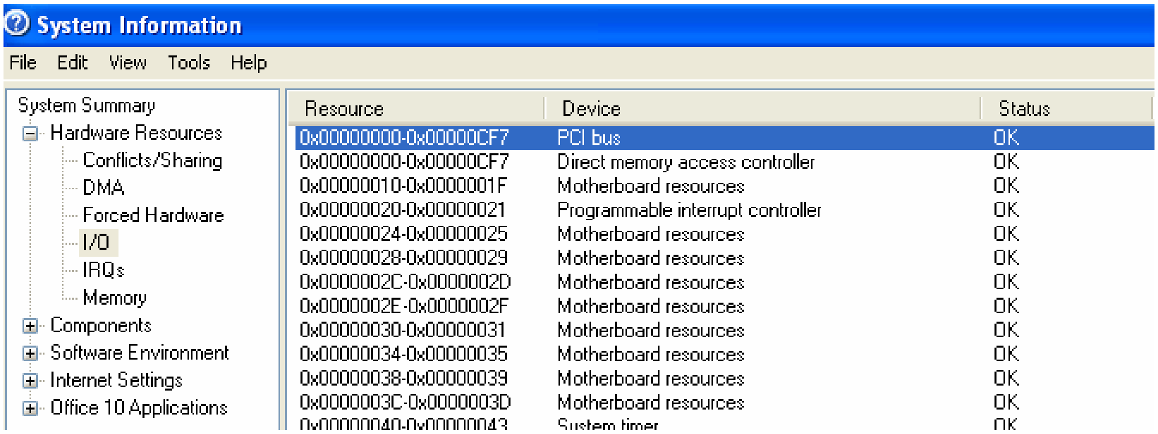
Note
There are several types of DMA in a modern PC. The DMA channels referred to in this section involve the ISA bus. Other buses, such as the ATA/IDE bus used by hard drives, have different DMA use. The DMA channels explained here don't involve your ATA/IDE drives, even if they are set to use DMA or Ultra DMA transfers.

I/O Port Addresses

I/O is located in the Hardware Resources category in System Information. I/O is the communication channel among hardware devices in your computer. The **Resource** column in the details pane displays the resource that is used by an I/O device, which appears in the **Device** column. The **Status** column displays the status of the device.

If there is a sharing conflict with a device that uses an I/O channel, locate the device in Conflicts/Sharing or in I/O, and then use Device Manager to locate and troubleshoot the problem.

The picture below shows some of the commonly used motherboard and chipset-based I/O port usage.



Motherboard and Chipset-Based Device Port Addresses

To find out exactly which port addresses are being used on your motherboard, click Start, Programs, Accessories, System Tools and System information. Then expand the Hardware resources so that you can see the IRQ information. Virtually all devices on your system buses use I/O port addresses. Most of these are fairly standardized, meaning you won't often have conflicts or problems with these settings. In the next section, you learn more about working with I/O addresses.

System Resource Conflict

In general, shared resources are not problematic, but investigating potential conflicts can help resolve issues. If you are having a problem with a device, check Conflicts/Sharing to verify that there are no sharing conflicts. Locate the device in the **Device** column in the details pane, and then verify which resource the device uses in the **Resource** column. Devices that share the same resource have duplicate entries in the **Resource** column. For example, you may have two adapters that share the same IRQ channel, such as IRQ 5. If the problem device shares a resource with another device, there may be a conflict. You can use Device Manager to troubleshoot the problem.

Conflicts/Sharing is located in the Hardware Resources category in System Information. It displays devices that share the same resources, such as direct memory access (DMA), IRQs, and memory.

The resources in a system are limited. Unfortunately, the demands on those resources seem to be unlimited. As you add more and more adapter cards to your system, you will find that the potential for resource conflicts increases. If your system is fully PnP-compatible, potential conflicts should be resolved automatically, but often are not.

How do you know whether you have a resource conflict? Typically, one of the devices in your system stops working. Resource conflicts can exhibit themselves in other ways, though. Any of the following events could be diagnosed as a resource conflict:
A device transfers data inaccurately.

- Your system frequently locks up.
- Your sound card doesn't sound quite right.
- Your mouse doesn't work.
- Garbage appears on your video screen for no apparent reason.
- Your printer prints gibberish.
- You cannot format a floppy disk.
- The PC starts in Safe mode (Windows 9x).

Windows 9x/Me and Windows 2000/XP also show conflicts by highlighting a device in yellow or red in the Device Manager representation. By using the Windows Device Manager, you can usually spot the conflicts quickly. In the following sections, you learn some of the steps you can take to head off resource conflicts or track them down when they occur.

Warning
Be careful in diagnosing resource conflicts; a problem might not be a resource conflict at all, but a computer virus. Many computer viruses are designed to exhibit themselves as glitches or periodic problems. If you suspect a resource conflict, it might be worthwhile to run a virus check first to ensure that the system is clean. This procedure could save you hours of work and frustration.

One way to resolve conflicts is to help prevent them in the first place. Especially if you are building up a new system, you can take several steps to avoid problems. One is to avoid using older ISA devices. By definition, they cannot share IRQs, and that is the resource most in demand. PCI (and AGP) cards can share IRQs with IRQ Steering and as such are a much better choice. Another way you can help is to install cards in a particular sequence and not all at once. Modifying the installation sequence often helps because many cards can use only one or two out of a predefined selection of IRQs that is specific to each brand or model of card. By installing the cards in a controlled sequence, the plug-and-play software can more easily work around IRQ conflicts caused by the default configurations of different cards.

The first time you start up a new system you have assembled or done major upgrades on, the first thing you should check is the BIOS Setup. If you have a setting for PnP Operating System in your BIOS, be sure it is enabled if you are running an operating system with plug-and-play support, such as Windows 9x/Me/2000. Otherwise, make sure it's disabled if you are running an OS that is not plug-and-play, such as Windows NT.

On initial startup I recommend a minimum configuration with only the graphics card, memory, and storage drives (floppy, hard disk, CD-ROM, and DVD). This allows for the least possibility of system conflicts in the initial configuration. If your motherboard came with a CD including drivers specific to the chipset or other built-in features of the board, now is the time to load or install them. Complete the configuration of all built-in devices before installing any other cards or external devices.

After the basic system has been configured (and after you have successfully loaded your operating system and any updates or patches), you can then begin adding one device at a time in a specific order. So, you will power down, install the new device, power up, and proceed to install any necessary drivers and configure the device. You'll probably have to restart your system after you are done to fully complete the configuration.

Following is the loading sequence for additional cards:

1. Sound card
2. Internal or external modem
3. Network card
4. Auxiliary video devices, such as MPEG decoders, 3D accelerators, and so on
5. SCSI adapter
6. Anything else

Normally, using this controlled sequence of configuring or building up your system results in easier integration with less conflicts and configuration hassles.

Resolving Conflicts Manually

In the past, the only way to resolve conflicts manually was to take the cover off your system and start changing switches or jumper settings on the adapter cards. Fortunately, this is a bit easier with plug and-play because the entire configuration is done via the Device Manager software included in the operating system. Although some early plug-and-play cards also had jumper switches or setup options to enable them to be configured manually, this feature was found primarily on ISA PnP-compatible cards. Be sure you write down or print out your current system settings before you start making changes. That way, you will know where you began and can go back to the original configuration (if necessary).

Finally, dig out the manuals for all your adapter boards; you might need them. Additionally, you could look for more current information online at the manufacturers' Web sites. Now you are ready to begin your detective work. As you try various resource settings, keep the following questions in mind; the answers will help you narrow down the conflict areas:

- *When did the conflict first become apparent?* If the conflict occurred after you installed a new adapter card, that new card probably is causing the conflict. If the conflict occurred after you started using new software, the software probably uses a device that is taxing your system's resources in a new way.
- *Are two similar devices in your systems that do not work?* For example, if your modem, integrated serial ports, or mouse—devices that use a COM port—do not work, chances are good that these devices are conflicting with each other.
- *Have other people had the same problem, and if so, how did they resolve it?* Public forums—such as those on CompuServe, Internet newsgroups, and America Online—are great places to find other users who might be able to help you solve the conflict.

Whenever you make changes in your system, reboot and see whether the problem persists. When you believe that you have solved the problem, be sure to test all your software. Fixing one problem often seems to cause another to crop up. The only way to ensure that all problems are resolved is to test everything in your system.
One of the best pieces of advice I can give you is to try changing one thing at a time, and then retest.

Common Resources Problems

Sound Cards

Sound cards are probably the biggest single resource hog in your system. They typically use at least one IRQ, two DMA channels, and multiple I/O port address ranges. This is because a sound card is actually several different pieces of hardware all on one board. Most sound cards, including PCI-based models, emulate the Sound Blaster 16 from Creative Labs.

Although other brands of sound cards might use a slightly different configuration, the pattern is the same; Sound Blaster emulation requires a large number of resources. Even if SB emulation isn't used (you might be able to disable it if you no longer play DOS games), the card still uses a single IRQ and several I/O port address ranges. If you take the time to read your soundboard's documentation and determine its communications-channel needs, compare those needs to the IRQ lines and DMA channels that already are in use in your system, and then change the settings of the other adapters to avoid conflicts with the sound card, your installation will go quickly and smoothly. Unfortunately, many vendors no longer provide detailed information on their plug-and-play–compatible cards. That's why you need to install the sound card **first** and use the system resource map to record which settings the card uses before you install other cards.

Recommendation
It is recommended that the sound card be installed before all other cards—except for video. In other words, let the sound card retain all its default settings if you can. Try to change the settings of other adapters when a conflict with the sound card arises. The problem here is that many older programs that use sound are very poorly written with respect to supporting alternative resource settings on sound cards. Save yourself some grief, and let the sound card have its way!
Even the latest sound cards can have problems if they're installed after other cards. I know a user who had to remove all his plug-and-play cards before his system would recognize his plug-and-play sound card.

Network Interface Cards

Networks are becoming more and more popular all the time, thanks to the rise of easy-to-configure small office/home office networks and the use of network cards to connect to broadband Internet devices such as cable and DSL modems. A typical network adapter does not require as many resources as some of the other cards discussed here, but it requires at least a range of I/O port addresses and an interrupt. Most network interface cards (NICs) also require a 16KB range of free upper memory to be used for the RAM transfer buffer on the network card. As with any other cards, be sure that all these resources are unique to the card and are not shared with any other devices.

Multiple-COM-Port Adapters

A serial port adapter usually has two or more ports onboard. These COM ports require an interrupt and a range of I/O ports each. There aren't too many problems with the I/O port addresses because the ranges used by up to four COM ports in a system are fairly well defined. The real problem is with the interrupts. Most older installations of more than two serial ports have any additional ones sharing the same interrupts as the first two. This is incorrect and causes nothing but problems with software that runs under Windows or OS/2. With these older boards, ensure that each serial port in your system has a unique I/O port address range and, more importantly, a unique interrupt setting.

Universal Serial Bus

USB ports are now found on most motherboards, and Windows 98, Windows Me, Windows 2000, and Windows XP provide you with a wide range of operating systems that support them properly. One potential problem is that USB takes another interrupt from your system, and many computers either don't have any free or are down to their last one. If your system supports PCI IRQ Steering, this shouldn't be much of a problem because the IRQ used by your USB controller should be sharable with other PCI devices.

If you are out of interrupts, you should look at what other devices you can disable (such as COM or LPT ports) to gain back a necessary interrupt for other devices.

The big advantage of USB from an IRQ or a resource perspective is that the USB bus uses only one IRQ no matter how many devices (up to 127) are attached. Therefore, you can freely add or remove devices from the USB without worrying about running out of resources or having resource conflicts.

Miscellaneous Boards

Some video cards ship with advanced software that allows special video features, such as oversized desktops, custom monitors, switch modes on-the-fly, and so on. Unfortunately, this software requires that the card be configured to use an IRQ. If your video card doesn't require an IRQ, I suggest you dispense with this unnecessary software and configure the card to free up the interrupt for other devices.

Plug-and-Play Systems

Plug and Play (PnP) represents a major revolution in recent interface technology. PnP first came on the market in 1995, and most motherboards and adapter cards since 1996 take advantage of it. Prior to that, PC users were forced to muddle through a nightmare of DIP switches and jumpers every time they wanted to add new devices to their systems. The results, all too often, were system resource conflicts and nonfunctioning cards.

PnP was not an entirely new concept. It was a key design feature of MCA and EISA interfaces that preceded it by almost 10 years, but the limited appeal of MCA and EISA meant that they never became true de facto industry standards. Therefore, mainstream PC users still had to worry about I/O addresses, DMA channels, and IRQ settings. Early PCI-based systems also used a form of PnP configuration, but because there was no provision for managing conflicts between PCI and ISA cards, many users still had configuration problems. But now that PnP has become prevalent, worry-free hardware setup is available to all computer buyers.

For PnP to work, the following components are desired:

- PnP hardware
- PnP BIOS
- PnP operating system

Each of these components needs to be PnP-compatible, meaning that it complies with the PnP specifications.

The Hardware Component

The *hardware component* refers to both computer systems and adapter cards. The term does not mean, however, that you cannot use your older ISA adapter cards (referred to as *legacy cards*) in a PnP system. You can use these cards; in fact, your PnP BIOS automatically reassigns PnP-compatible cards around existing legacy components. Also, many late-model ISA cards can be switched into PnP-compatible mode.

PnP adapter cards communicate with the system BIOS and the operating system to convey information about which system resources are necessary. The BIOS and operating system, in turn, resolve conflicts (wherever possible) and inform the adapter cards which specific resources it should use. The adapter card then can modify its configuration to use the specified resources.

The BIOS Component

The BIOS component means that most users of pre-1996 PCs need to update their BIOSes or purchase new machines that have PnP BIOSes. For a BIOS to be compatible, it must support 13 additional system function calls, which can be used by the OS component of a PnP system. The PnP BIOS specification was developed jointly by Compaq, Intel, and Phoenix Technologies.
The PnP features of the BIOS are implemented through an expanded POST. The BIOS is responsible for identification, isolation, and possible configuration of PnP adapter cards. The BIOS accomplishes these tasks by performing the following steps:

1. Disables any configurable devices on the motherboard or on adapter cards
2. Identifies any PnP PCI or ISA devices
3. Compiles an initial resource-allocation map for ports, IRQs, DMAs, and memory
4. Enables I/O devices
5. Scans the ROMs of ISA devices
6. Configures initial program-load (IPL) devices, which are used later to boot the system
7. Enables configurable devices by informing them which resources have been assigned to them
8. Starts the bootstrap loader
9. Transfers control to the operating system

The Operating System Component

The operating system component is found in most modern operating systems, such as Windows 9x/Me/2000/XP. In some cases system manufacturers have provided extensions to the operating system for their specific hardware. Such is especially true for notebook

systems, for example. Be sure you load these extensions if they are required by your system.

It is the responsibility of the operating system to inform users of conflicts that cannot be resolved by the BIOS. Depending on the sophistication of the operating system, the user then could configure the offending cards manually (onscreen) or turn off the system and set switches on the physical cards. When the system is restarted, the system is checked for remaining (or new) conflicts, any of which are brought to the user's attention. Through this repetitive process, all system conflicts are resolved.

Note
Because of revisions in some of the Plug and Play specifications, especially including the ACPI specification, it can help to ensure you are running the latest BIOS and drivers for your system. With the Flash ROM used in most PnP systems, you can download the new BIOS image from the system vendor or manufacturer and run the supplied BIOS update program.

Hardware Selection Criteria

A system might not have to meet every one of these criteria for you to consider purchasing it, but if it misses more than a few, consider staying away from that system. The items at the top of the list are the most important, and the items at the bottom are perhaps of lesser importance (although I think each item is important). The rest of this chapter discusses in detail the criteria in this checklist:

- Processor. A modern system should use a socket-based processor with on-die L2 cache. Evaluate the processor choices you have, and try to get the one with the highest speed CPU bus (frontside bus). Don't get too hung up on L2 cache size; a little cache goes a long way. It is more important that the cache run at full core speed (which it will if it is on-die). Processors such as the later Celeron/Pentium III, Duron/Athlon, and Pentium 4 all meet this criteria.
- Processor sockets. For maximum upgradability and performance, you should stick with a system that uses a socket for the CPU. The main sockets in use today are Socket 370 for the Celeron/Pentium III, Socket A (Socket 462) for the Duron/Athlon, and Socket 423 for the Pentium 4. As long as your motherboard has one of these sockets, you should be in good shape. Motherboard speed. The motherboard typically offers a choice of speeds, including anywhere from 66MHz to 266MHz for Celeron/Pentium III–based boards, 200MHZ–266MHz for the Duron/Athlon-based boards, or 400MHz for the Pentium 4–based boards. Check to ensure the board you are buying runs at the speeds necessary to support the processors you want to install.

- Cache memory. All modern systems use processors with integral cache, most of them now having the cache directly on the processor die for maximum speed. As such, there won't be any cache memory on the motherboard in a modern system. The tip is to make sure you are using a processor with full core speed on-die L2 cache because this offers the maximum in performance. All the modern processors now incorporate on-die L2 cache.
- SIMM/DIMM/RIMM memory. SIMMs are obsolete by today's standards, so stay away from any boards that use them. Your board should support either standard or DDR DIMMs or RIMMs that contain SDRAM, DDR SDRAM, or RDRAM, respectively. What you use depends mainly on your motherboard chipset, so choose the chipset and board that accepts the memory type you want to use. Currently, DDR SDRAM and RDRAM are the fastest types of memory available, with RDRAM being by far the most costly. Mission-critical systems should use ECC memory and ensure that the motherboard fully supports ECC operation. Note that many of the low-end chipsets, such as the Intel 810/815 series, do not support ECC and should not be used for mission-critical applications. This is something you should know before purchasing the system. Finally, note that most motherboards support either three or four DIMM sockets, or two or three RIMM sockets. Be sure that you populate them wisely so you don't have to resort to removing memory later to add more, which is not very cost-effective.
- Bus type. Most motherboards no longer incorporate ISA bus slots, so if you have any ISA cards, you'll likely have to discard them. Instead you should find anywhere from one to five or more PCI local bus slots. Be sure the PCI slots conform to the PCI 2.1 or later revision (primarily based on the chipset). Take a look at the layout of the slots to ensure that cards inserted in them will not block access to memory sockets or be blocked by other components in the case. Systems without onboard video should also feature one AGP 4x slot for a high-performance AGP video card.
- BIOS. The motherboard should use an industry-standard BIOS, such as those from AMI, Phoenix, or Award. The BIOS should be of a Flash ROM or EEPROM design for easy updating. Look for a BIOS Recover jumper or mode setting, as well as possibly a Flash ROM write-protect jumper on some systems.
- Form factor. For maximum flexibility, performance, reliability, and ease-of-use, the ATX form factor (including micro-ATX and flex-ATX) cannot be beat. ATX has several distinct performance and functional advantages over Baby-AT and is vastly superior to any proprietary designs, such as LPX. Additionally, the new NLX form factor might be a consideration for low-profile or low-cost desktop systems.
- Built-in interfaces. Ideally, a motherboard should contain as many built-in standard controllers and interfaces as possible (except perhaps video). There is a trend toward legacy-free PCs that lack the conventional Super I/O component and therefore have only USB for external expansion. Legacy-free PCs lack the conventional keyboard and mouse ports, serial and parallel ports, and possibly

even the internal floppy controller. Systems that use an integrated Super I/O component have these interfaces. Built-in 10/100 Ethernet network adapters are also handy, especially if you are using a cable modem or DSL connection to the Internet. A built-in sound card is a great feature, usually offering full Sound Blaster compatibility and functions, and possibly offering additional features as well. If your sound needs are more demanding, you might find the built-in solutions less desirable, and you might want to have a separate sound card in your system. Built-in video adapters are also a bonus in some situations, but because there are many video chipset and adapter designs from which to choose, generally you'll find better choices in external local bus video adapters. This is especially true if you need the highest performance video available. Built-in devices usually can be disabled to allow future add-ons, but problems can result.

- Onboard IDE interfaces. All motherboards on the market have included onboard IDE interfaces for some time now, but not all IDE interfaces are equal. Your motherboard should support at least UDMA/33 (ATA-33) speeds, which matches the best real-world performance currently available from IDE drives. UDMA/66 and UDMA/100 actually exceed the real-world performance of current drives using these standards and thus provide you with headroom for future drives. For even greater speed, consider motherboards with onboard IDE RAID controllers. These motherboards can be configured to perform data striping for extra speed or data mirroring for extra reliability when two or more identical IDE drives are used. These motherboards are based on a variety of standard chipsets with RAID functions added by RAID chipsets from AMI, HighPoint, or Promise.
- Power management. The motherboard should fully support the latest standard for power management, which is ACPI. An Energy Star–compliant system is also a bonus because it uses less than 30 watts of electrical energy when in sleep mode, saving energy as well as your electric bill.
- Motherboard chipset. Motherboards should use a high-performance chipset that supports standard or DDR SDRAM DIMMs or RDRAM RIMMs—preferably one that allows ECC memory as well, if you are concerned about catching possible memory errors before they corrupt your data. Also look for AGP 4x or faster video support and ATA-100 hard drive support.
- Documentation. Good technical documentation is a requirement. Documents should include information on any and all jumpers and switches found on the board, connector pinouts for all connectors, specifications for other plug-in components, and any other applicable technical information. I would also acquire separate documentation from the BIOS manufacturer covering the specific BIOS used in the system, as well as the data books covering the specific chipset used in the motherboard. Additional data books for any other controller or I/O chips onboard are a bonus and can be acquired from the respective chip manufacturers. Another nice thing to have is available online support and documentation updates, although this should not be accepted in place of good hardcopy manuals. Remember that many vendors provide their "hardcopy" manuals as Adobe

Acrobat files, which can be viewed and printed with the Adobe Acrobat Reader program.

You might notice that these selection criteria seem fairly strict and might disqualify many motherboards on the market, including what you already have in your system! These criteria will, however, guarantee you the highest-quality motherboard offering the latest in PC technology that will be upgradable, be expandable, and provide good service for many years. Most of the time I recommend purchasing boards from better-known motherboard manufacturers such as Intel, Acer, ABIT, AsusTek, SuperMicro, Tyan, and others. These boards might cost a little more, but there is some safety in the more well-known brands. That is, the more boards they sell, the more likely that any problems will have been discovered by others and solved long before you get yours. Also, if service or support is necessary, the larger vendors are more likely to be around in the long run.

Warning

Most chipset manufacturers make a particular chip for only a short time, rapidly superseding it with an improved or changed version. The data books are available only during the time the chip is being manufactured, so if you wait too long, you will find that such documents might no longer be available. Get the documentation for your motherboard as soon as possible after buying the motherboard.

ROM BIOS

ROM BIOS

ROM (read-only memory) is a type of memory that stores data even when the main computer power is off. This is necessary so that the system can access the data it needs to start up. When stored in ROM, information that is required to start and run the computer cannot be lost or changed. The *BIOS*, software in the form of programs stored on ROM chips, is used during the startup routine to check out the system and prepare to run the hardware. The *system BIOS* is a ROM chip on the motherboard used by the computer during the startup routine (boot process) to check out the system and prepare to run the hardware. The BIOS is stored on a ROM chip because ROM retains information even when no power is being supplied to the computer. The downside of storing data in an older computer's ROM is that we have to change a chip to update information.

More recent systems use a technology called flash ROM or flash BIOS that allows code in the core chips to be updated by software available through the BIOS or motherboard supplier. Check the Internet site for the supplier if you suspect your ROM chip has flash ROM technology; the software and instructions are generally downloadable.

Upgrade a BIOS only when necessary! Be sure to follow all precautions included with the motherboard manual and instructions for the upgrade. Improper installation can render the motherboard useless.

BIOS (also referred to as firmware) can be subdivided into three classes, depending upon the type of hardware it controls.

- The first class, called core chips, includes support for hardware that is common to all computers, is necessary, and never changes.
- The second class, called updatable chips, encompasses hardware that is also common and necessary, but that might change from time to time.
- The third class of chips includes anything that is not included in one of the first two classes.

Installing BIOS

If you have an old computer that can't play with the newer parts available, one option is to upgrade your BIOS. The BIOS in some computers just can't handle things like CD-ROMs or even 3½" floppy drives. BIOS won't make your computer run any faster; it just helps it play with the big boys in a slow manner. Finding the BIOS chips can be somewhat frustrating. To make sure you get the right kind, you can go to the place where

you bought your motherboard and see if they have any left. You can also look in computer magazines for ads that advertise ROM-BIOS. Just make sure you know what kind you're looking for. (A description as to how to find your BIOS is on the Motherboards page.) You can also find out what type you have on start-up; watch the screen, and it may say something similar to "AMIBIOS ©1992."

Installing the BIOS

1. Turn off the computer, unplug it, and take the case off.
2. Find your old BIOS. If you have several BIOS chips, note their order--sometimes they are numbered. The new BIOS will need to be put in in the same order.
3. Remove the old chips. If there is anything in the way, get it out. Now, get a chip-puller, a tweezer-like tool. This makes it easier to yank your BIOS chips out. If you don't have a chip-puller, you can gently use a flat-head screwdriver to lift up each side of the chip little by little.
4. Insert the new ones in the correct order. Also, make sure the notched end of the chip matches the notched end of the socket. Make sure that the little pins on the chip are straight. While installing them, put the first row of pins in first. Then slowly lower the other side into the socket, making sure they line up with the holes. After you're finished, give the chips a final push down, just to make sure.
5. Put everything you took out back in. Clean up your mess. Put the case on.
6. Turn your computer on. It should boot up and give you a new BIOS date on the screen.

A Typical CMOS Setup

Every CMOS setup program looks slightly different. Do not be too concerned about the differences—all BIOS routines contain basically the same information. Take your time and be comfortable searching around the setup programs. Most of the CMOS setup programs are text-based, so you will have to use keystrokes to navigate through the information. However, some newer machines use a Windows-like CMOS setup (they have the look of a Windows environment and will let you use a mouse to select changes).

The CMOS Battery

The CMOS chip requires a small trickle voltage from a battery to keep its memory alive. When the battery gets low or runs out of current, the computer will experience a sudden memory loss and thus lose settings.

The voltage of CMOS batteries ranges from 3 to 6 volts. Check the motherboard or the motherboard documentation to determine the actual battery requirements. Batteries come as either on-board (NiCad batteries, soldered in place or in a fixture, that last from five to

seven years) or external (no rechargeable AA alkaline batteries that last from two to four years).

The first clue that the battery is weakening is that the CMOS clock begins to slow down. Go to a C prompt and type **time.** If you notice the clock is slow, it's time to change the battery.

Remember that an MS-DOS machine (this includes one running Windows 95 or 98) uses the CMOS clock only to get the date upon startup. After the computer is running, MS-DOS uses the memory refresh timer on the memory controller to keep time. This works well, but because the refresh timer is not very accurate about seconds, you will lose one or two seconds per day. If you never turn off a computer, it could lose time. Do not confuse this with a bad battery. When you reboot, the computer will update itself to the correct time from the CMOS. If the CMOS battery is low, it will still show the incorrect time.

When the CMOS battery dies completely, you will get lost CMOS errors, as previously described. If you reload the CMOS data and the errors return, you must change the battery. Although the computer will hold CMOS information during the week, sometimes, over the weekend—when the computer is turned off for two days—the CMOS data will be lost. Do not let these seemingly "intermittent" problems fool you. And sometimes, before a battery dies, but after it has started to fail, it will still be able to hold CMOS settings for a short time after the computer is off. Any time a computer loses the CMOS information more than once in a week, replace the battery immediately—if only to eliminate the battery as the source of the problem. After replacing the battery, you must run the setup utility and restore any CMOS settings.

ROM Chips with BIOS

BIOS can be put on the hardware device itself. Many new add-on boards such as display adapters, network interfaces, and sound cards have their own on-board ROM chip. Since the system BIOS doesn't have a clue about how to talk to the device, this card brings its own BIOS.

Loading Device Drivers

Did you ever wonder why (almost) every newly purchased device comes with an installation disk? Did you ever notice the line(s) DEVICE= in the CONFIG.SYS file (a text file found in the root directory of the hard disk drive) or how often a program writes information to Windows 98 or 2000 Registry?

Using device drivers is the most popular way to provide BIOS support for hardware. A device driver is a program that acts as an interface between the operating system and the

control circuits that operate the device. For example, Windows has "generic" code that opens a file, but the driver for the disk drive takes care of low-level tasks like positioning the read head, reading or writing blocks of data, and so on. Thus, applications programmers don't usually have to worry about these details and can assume that any hardware supported by a device driver will work. Every time the computer is booted up, the CONFIG.SYS file is read and the device drivers are loaded from the hard disk drive into RAM.

Some examples of device drivers in CONFIG.SYS are:

```
DEVICE=C:\DOS\HIMEM.SYS
DEVICE=C:\DOS\EMM386.EXE NOEMS
DEVICEHIGH =C:\SCSI\ASPIPP3.SYS /D /Z
DEVICEHIGH =C:\DOS\DRVSPACE.SYS
DEVICEHIGH =C:\CDROM\MTMCDAI.SYS /L=001
```

Loading device drivers in CONFIG.SYS is a requirement for machines running MS-DOS. The Windows 95, 98, and 2000 operating systems have their own drivers that are loaded as part of startup. (Drivers are covered in more detail in Chapter 15, "Software: MS-DOS and Windows 3.x," and Chapter 16, "Windows 98 and Beyond.") Occasionally, drivers become outdated or have problems. You can obtain new drivers directly from the device manufacturer (frequently from their Web sites).

Even hardware that installs without a setup disk can be changing the registry if it is a Plug and Play device that is recognized by the operating system. Erratic problems can occur if a device is improperly identified. Under Windows 95, 98, or 2000, check under the System section of Control Panel to identify possible conflicts.

POST—Power-On Self Test

Every time a PC is turned on or reset using the Reset button or Windows Restart command, the computer is rebooted and reset to its basic operating condition. The system BIOS program starts by invoking a special program (stored on a ROM chip) called the *POST (power-on self test)*. The POST sends out standardized commands that check every primary device (in more technical terms, it runs an internal self-diagnostic routine).

The POST has two stages:

- Test 1 occurs before and during the test of the video.
- Test 2 occurs after the video has been tested.

This division determines whether the computer will display errors by beeping or showing them on the screen. The POST does not assume the video works until it has been tested. The POST does assume that the speaker always works, but in order to let you know that

the speaker is working, all computers beep on startup. Depending on the BIOS type, the POST might also sound a single beep when it's done, to let you know the boot process was successful. If something goes wrong, the POST sends a series of beep codes to let you know what the problem is or where to start looking for it.

Beep Codes Before and During the Video Test

The purpose of the first POST test is to check the most basic components. The exact order, number of tests, and error states will vary from product to product. In a healthy system, the POST reports by using a series of beep codes and screen messages to convey that all components are working. Then it transfers control to the boot drive, and the operating system is loaded. The POST is a good indication that the hardware is in working order.

If a problem occurs, the POST routine attempts to report the problem. This is also done by beep codes and (if possible) screen prompts. Some error codes are specific to chip sets or custom products, and the exact message and its meaning can vary from system to system. (See the POST code references in the system manual that shipped with the PC or the motherboard to obtain references for detailed error messages and beeps.) The following table lists the basic beep codes for AMI and Phoenix BIOSs.

Number of Beeps	**Possible Problem**
1	DRAM refresh failure
2	Parity circuit failure
3	Base 64 KB or CMOS RAM failure
4	System timer
5	Processor failure
6	Keyboard controller or Gate A20 error
7	Virtual mode exception error
8	Display monitor write/read test failure
9	ROM BIOS checksum error
10	CMOS RAM shutdown register failure

1 long, 3 short	Conventional/extended memory test failure
1 long, 8 short	Display test and display vertical and horizontal retrace test failure

Troubleshooting After a Beep

After a beep code has been recognized, there are a few things you can do to troubleshoot the error. The following table suggests some solutions. Keep in mind that in many cases it can be less expensive to replace the motherboard than to replace a chip.

Problem	Solution
RAM refresh failure Parity error RAM bit error Base 64-KB error	Reseat and clean the RAM chips. Replace individual memory chips until the problem is corrected.
8042 error (keyboard chip) Gate A20 error	Reseat and clean keyboard chip. Check operating system. Replace keyboard. Replace motherboard.
BIOS checksum error	Reseat ROM chip. Replace BIOS chip.
Video errors	Reseat video card. Replace video card.
Cache memory error	Reseat and clean cache chips. Verify cache jumper settings are correct. Replace cache chips.
Any other problems	Reseat expansion cards. Clean motherboard. Replace motherboard.

NOTE

Many computers will generate beep codes when the only problem is a bad power supply! Turn the computer off and on three or four times to see if the same beep code is generated every time. If so, it's probably a legitimate beep code that concerns the hardware and not the power supply.

After successfully testing the video, the POST will display any error messages on the screen.

BIOS may consist of three parts: -

- **Motherboard ROM BIOS**
  The main part of the BIOS is burned or flashed into a ROM chip that is nonvolatile (it is not erased when the power is turned off) and read-only.

- **Adaptor Card ROM**
  ROM chips installed on adapter cards

- **Disk Drivers**
  Additional drivers loaded when your system boots up.

The combination of the above three components contribute to the BIOS as a whole.

Motherboard ROM BIOS Functions

All motherboards have a special chip containing software called the ROM BIOS. This ROM chip contains the startup programs and drivers used to get the system running and act as the interface to the basic hardware in the system.
BIOS in most PCs has four main functions: -

POST (power on self test).

The POST tests your computer’s processor, memory, chipset, video adapter, disk controllers, disk drives, keyboard, and other crucial components.

Setup.

System configuration and setup program. This is usually a menu-driven program activated by pressing a special key during the POST, which enables you to configure the motherboard and chipset settings along with the date and time, passwords, disk drives, and other basic system settings. You also can control the power-management settings and boot-drive sequence from the BIOS Setup, and on some systems, you can also configure CPU timing and clock-multiplier settings.

ROM Hardware

Read-only memory (ROM) is a type of memory that can permanently or semi-permanently hold data. It is called read-only because it is difficult to write to. ROM is also often called nonvolatile memory because any data stored in ROM remains even if

the power is turned off. This is why ROM is a best place to put the PC's startup instructions. Note that ROM and RAM are not opposites. In fact, ROM is technically a subset of the system's RAM. In other words, a portion of the system's random access memory address space is mapped into one or more ROM chips. This is necessary to contain the software that enables the PC to boot up; otherwise, the processor would have no program in memory to execute when it is powered on.

Adapter card ROMs, such as those used for video, SCSI, add-on IDE cards, and network cards for diskless workstations, are automatically scanned and read by the motherboard ROM during the early part of the boot process, during the POST. The motherboard BIOS automatically runs the programs in any adapter ROMs it finds during the scan. You see this in most systems when you turn on your system, and during the POST you see the video card BIOS initialize and announce its presence.

BIOS Upgrade and Recovery

ROM BIOS upgrades can improve a system in many ways. The ROM BIOS provides intelligence that integrates the working of the components of a computer. BIOS communicates with the hardware and it must be specific to the hardware. Often, in older systems, you must upgrade the BIOS to take advantage of some other upgrade. To install some of the larger and faster IDE, for example, you might need a BIOS upgrade.

BIOS Version

You must know the following information before you upgrade a BIOS for a particular motherboard.

- The make and model of the motherboard (or system)
- The version of the existing BIOS
- The type of CPU (for example, Pentium MMX, AMD K6, Cyrix/IBM 6x86MX, MII, Pentium II, Pentium III and later, AMD Athlon, and so on)

In the older systems you can identify the BIOS by watching the screen when the system started up. BIOS information is displayed for only a few seconds. In addition, you often can find the BIOS ID information in the BIOS Setup screens. You usually can press the Pause key on the keyboard when the BIOS ID information is being displayed; this freezes it so you can record the information. Pressing any other key allows the system startup to resume. Unicore Software also offers a downloadable BIOS Agent that can be used to determine this information, as well as the motherboard chipset and Super I/O chip used by your motherboard.
After you have this information, you should be able to contact the motherboard manufacturer to see whether a new BIOS is available for your system. If you go to the

Web site, check to see whether a version exists that is newer than the one you have. If so, you can download it and install it in your system.

Backing Up Your BIOS's CMOS Settings

A motherboard BIOS upgrade usually wipes out the BIOS Setup settings in the CMOS RAM. Therefore, you should record these settings, especially the important ones such as hard disk parameters. Manually record your BIOS Setup parameters, or possibly connect a printer to your system and use the Shift+Prtsc (Print Screen) function to print out each of the setup screens.

Using a Flash BIOS

Virtually all PCs built since 1996 include a flash ROM to store the BIOS. A flash ROM is a type of EEPROM chip you can erase and reprogram directly in the system without special equipment. Using flash ROM enables you to download ROM upgrades from a Web site or to receive them on disk; you then can load the upgrade into the flash ROM chip on the motherboard without removing and replacing the chip. Normally, these upgrades are downloaded from the manufacturer's Web site, and then an included utility is used to create a bootable floppy with the new BIOS image and the update program. It is important to run this procedure from a boot floppy so that no other software or drivers are in the way that might interfere with the update. This method saves time and money for both the system manufacturer and end user. Here is a step-by-step procedure for the process using Intel's iFlash (DOS-based) BIOS update:

1. Save your CMOS RAM setup configuration. You can do so by pressing the appropriate key during boot to start the BIOS setup program (usually F1 with an AMI BIOS and F2 with a Phoenix BIOS) and writing down all your current CMOS settings. You also might be able to print the screens if you have a printer connected, using the PrtScr key on the keyboard. You must reset these settings after you have upgraded to the latest BIOS. Write down all settings that are unique to the system. These settings will be needed later to reconfigure the system. Pay special attention to any hard drive settings for geometry (Cylinder/Head/Sectors per track) and translation (LBA, Large, CHS); these are very important. If you fail to restore these properly, you might not be able to boot from the drive or access the data on it.
2. Exit the BIOS Setup and restart the system. Allow the system to fully start Windows and bring up a DOS prompt window or boot directly to a DOS prompt via the Windows Start menu (for example, press F8 when you see Starting Windows, and select Command Prompt).
3. Place a formatted blank floppy disk in the A: floppy drive. If the disk contains data, format the floppy using the normal format command:
C:\>FORMAT A:
You also can use Windows Explorer to format the floppy disk.

4. The file you originally downloaded from the Intel Web site is a self-extracting compressed archive that includes other files that need to be extracted. Put the file in a temporary directory, and then from within this directory, double-click the BIOS file you downloaded or type the filename of the file and press Enter. This causes the file to self-extract. For example, if the file you downloaded is called CB-P06.EXE (for the Intel D810E2CB motherboard), you would enter the following command:
C:\TEMP>CB-P06 <enter>
5. The extracted files are stored in the same temporary folder as the downloaded BIOS. Recent Intel flash BIOS upgrades contain the following files: Desc.txt, License.txt, Readme.txt, Run.bat (which is run to create the bootable floppy), and SW.EXE (which contains the BIOS code).
6. To create the bootable floppy disk, open Run.bat, which extracts files from SW.EXEs and transfer the necessary files to the blank disk in Drive A:.
7. Now you can restart the system with the bootable floppy in drive A: containing the new BIOS files you just extracted. Upon booting from this disk, the iFLASH program automatically starts and updates the BIOS boot block and the main BIOS area.
8. When you're told that the BIOS has been successfully loaded, remove the bootable floppy from the drive and press Enter to reboot the system.
9. Press F1 or F2 to enter Setup. On the first screen within Setup, check the BIOS version to ensure that it is the new version.
10. In Setup, load the default values. If you have an AMI BIOS, press the F5 key. With a Phoenix BIOS, go to the Exit submenu and highlight Load Setup Defaults, and then press Enter.
11. If the system had unique settings, reenter those settings now. Press F10 to save the values, exit Setup, and restart the system. Your system should now be fully functional with the new BIOS.
Note
If you encounter a CMOS checksum error or other problems after rebooting, try rebooting the system again. CMOS checksum errors require that you enter Setup, check and save your settings, and exit Setup a second time.

Note

The procedure for older Intel BIOS upgrades differs from that described previously. If your BIOS upgrade contains the
BIOS.EXE program, see the BIOS upgrade description contained in *Upgrading and Repairing PCs,12th Edition*, included in electronic form on the CD-ROM packaged with this book.

Replacing a BIOS ROM

Systems dating from 1995 or earlier usually don't have a flash ROM and instead use an EPROM. To upgrade the BIOS in one of these systems, you replace the EPROM chip

with a new one preloaded with the new BIOS. As with a flash ROM upgrade, you must get this from your motherboard manufacturer.
The procedure for replacing the BIOS chip is also useful if you have made a backup copy of your socketed system BIOS chip and need to replace a damaged original with the backup copy.

To replace the BIOS chip, follow these steps:
1. Back up the CMOS RAM settings.
2. Power down the system and unplug the power cord.
3. Remove the cover and any other components in the way of the BIOS EPROM chip. Remember to use caution with respect to static discharges; you should wear an antistatic wrist strap for this procedure or ground yourself to the chassis before touching any internal components.
4. Using a chip puller or a thin flat-blade screwdriver, gently pry the chip out of its socket.
5. Remove the new EPROM from the antistatic packing material in which it came.
6. Install the new EPROM chip into the socket.
7. Reinstall anything you removed to gain access to the chip.
8. Put the cover back on, plug in the system, and power on.
9. Enter the BIOS setup information you saved earlier.
10. Reboot and enjoy the new BIOS.

IML System Partition BIOS

IBM and Compaq used a scheme similar to a flash ROM, called Initial Microcode Load (IML), in some of their older Pentium and 486 systems. IML is a technique in which the BIOS code is installed on the hard disk in a special hidden-system partition and is loaded every time the system is powered up. Of course, the system still has a core BIOS on the motherboard, but all that BIOS does is locate and load updated BIOS code from the system partition. This technique enabled Compaq and IBM to distribute ROM updates on disk for installation in the system partition. The IML BIOS is loaded every time the system is reset or powered on.

Chapter 7

Memory

OBJECTIVES

Memory (RAM)

Memory

Memory is the workspace for the computer's processor. It is a temporary storage area where the programs and data being operated on by the processor must reside. Memory storage is considered temporary because the data and programs remain there only as long as the computer has electrical power or is not reset. Before being shut down or reset, any data that has been changed should be saved to a more permanent storage device of some type (usually a hard disk) so it can be reloaded into memory again in the future.

RAM stands for Random Access Memory. Main memory is called RAM because you can randomly access any location in memory. This designation is not accurate. For example Read-only memory (ROM), is also randomly accessible, yet usually is differentiated from the system RAM because it maintains data without power and can't normally be written to. The disk memory is also randomly accessible. RAM is usually constructed of a type of chip called Dynamic RAM (DRAM).

DRAM (Dynamic RAM)

DRAM stores data dynamically. The information can be written to RAM again and again at any time. DRAM requires the data to be refreshed (essentially rewritten) every 15ms (milliseconds) or so. A type of RAM called Static RAM (SRAM) does not require the periodic refreshing. An important characteristic of RAM in general is that data is stored only as long as the memory has electrical power.

The main advantages of DRAM are that it is very dense, meaning you can pack a lot of bits into a very small chip. DRAM uses only one transistor and capacitor pair per bit, which makes them very dense. It is inexpensive compared to many other types of memory. The memory cells in a DRAM chip are tiny capacitors that retain a charge to indicate a bit. The problem with DRAM is that it is dynamic. Also, because of the design, it must be constantly refreshed; otherwise, the electrical charges in the individual memory capacitors will drain and the data will be lost.

Types of Computer Memory

Three main types of physical memory are used in modern PCs:

- ROM. Read-only memory
- DRAM. Dynamic random access memory
- SRAM. Static RAM

The speed of memory used to be expressed in ns (nanoseconds), and that of processor in MHz (megahertz). Newer and faster types of memory have speeds expressed in MHz. A *nanosecond* is defined as one billionth of a second.

Fast Page Mode DRAM

Standard DRAM is accessed through a technique called *paging*. Normal memory access requires that a row and column address be selected, which takes time. Paging enables faster access to all the data within a given row of memory by keeping the row address the same and changing only the column. Memory that uses this technique is called *Page Mode* or *Fast Page Mode* memory.

Extended Data out RAM

EDO memory consists of specially manufactured chips that allow a timing overlap between successive accesses. The effect of EDO is that cycle times are improved by enabling the memory controller to begin a new column address instruction while it is reading data at the current address.

SDRAM

Synchronous DRAM (SDRAM) is for Synchronous DRAM, a type of DRAM that runs in synchronization with the memory bus. SDRAM delivers information in very high-speed bursts using a high-speed, clocked interface. SDRAM removes most of the latency involved in asynchronous DRAM because the signals are already in synchronization with the motherboard clock. SDRAM is available in DIMM form and often is rated by megahertz speed rather than nanosecond cycling time.

To meet the stringent timing demands of its chipsets, Intel created specifications for SDRAM called PC66, PC100, and PC133. To meet the PC100 specification, 8ns chips usually are required. Normally, you would think 10ns would be considered the proper rating for 100MHz operation, but the PC100 specification calls for faster memory to ensure all timing parameters are met. The table shows the timing, actual speed, and rated speed for various SDRAM DIMMs.

SDRAM Timing, Actual Speed, and Rated Speed

Timing	Actual Speed	Rated Speed
15ns	66MHz	PC66
10ns	100MHz	PC66
8ns	125MHz	PC100
7.5ns	133MHz	PC133

DDR SDRAM

Double data rate (DDR) SDRAM memory is an evolutionary design of standard SDRAM in which data is transferred twice as quickly. Instead of doubling the actual clock rate,

DDR memory achieves the doubling in performance by transferring twice per transfer cycle: once at the leading (falling) edge and once at the trailing (rising) edge of the cycle. This is similar to the way RDRAM operates and effectively doubles the transfer rate, even though the same overall clock and timing signals are used. DDR SDRAM uses a new DIMM module design with 184 pins. Figure 8.2 shows the DDR SDRAM DIMM.

DDR DIMMs are rated for either PC1600 (200MHz x 8) or PC2100 (266MHz x 8) operation and normally run on 2.5 volts. They are basically an extension of the PC100 and PC133 DIMMs redesigned to support double clocking, where data is sent on each clock transition (twice per cycle) rather than once per cycle as with standard SDRAM.

RDRAM

RDRAM is a wide-channel device. They transfer data only 16 bits (2 bytes) at a time (plus 2 optional parity bits), but at much faster speeds. RIMM typically runs at 800MHz, so the overall throughput is 800 × 2, or 1.6GB per second for a single channel—the same as PC1600 DDR SDRAM. Dual-channel PC800 RDRAM is also currently the fastest type of PC memory, followed by PC2100 DDR SDRAM. Whether your board supports (requires) dual-channel RDRAM is dictated by the motherboard chipset; currently, the Intel 840 and 850 chipsets support dual-channel RDRAM.

The RIMMs themselves as well as the RDRAM devices run on only 2.5 volts. RDRAM chips are installed in modules called RIMMs. A RIMM similar in size and physical form to current DIMMs, but they are not interchangeable. RIMMs are available in 64MB, 128MB, and 256MB capacities. They can be added to a system one at a time because each individual RIMM technically represents multiple banks to a system.

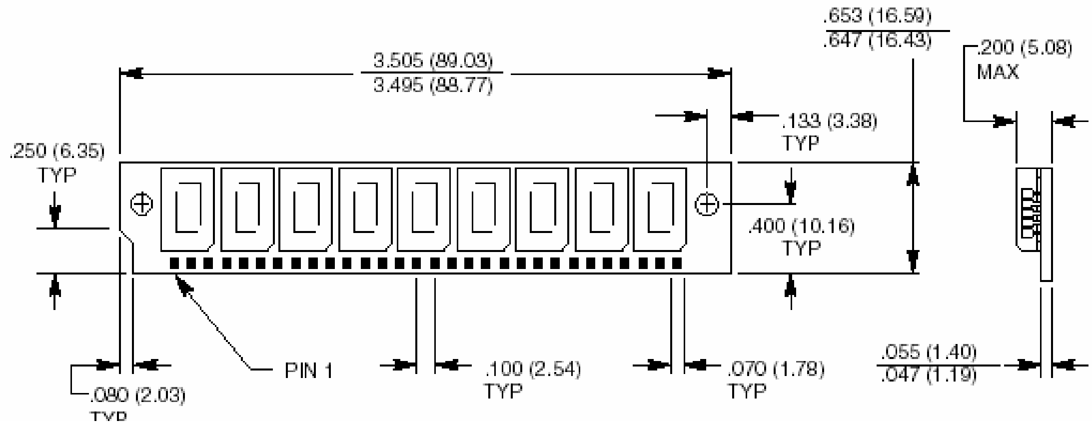
SIMMs, DIMMs, and RIMMs

Originally, systems had memory installed via individual chips. They are often referred to as *dual-inline-package (DIP)* chips because of their designs. The original IBM XT and AT had 36 sockets on the motherboard for these individual chips.
There are two main types of SIMMs, two main types of DIMMs, and so far one type of RIMM used in desktop systems. The various types are often described by their pin count, memory row width, or memory type.

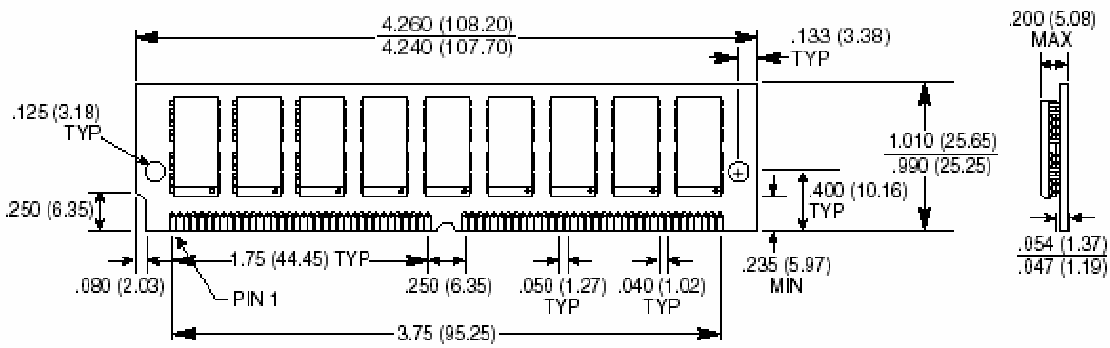
SIMM (Single Inline Memory Module)

SIMMs were available in two main physical types:
30-pin (8 bits plus an option for 1 additional parity bit)
72-pin (32 bits plus an option for 4 additional parity bits)

The 30-pin SIMM is physically smaller than the 72-pin versions, and either version can have chips on one or both sides.



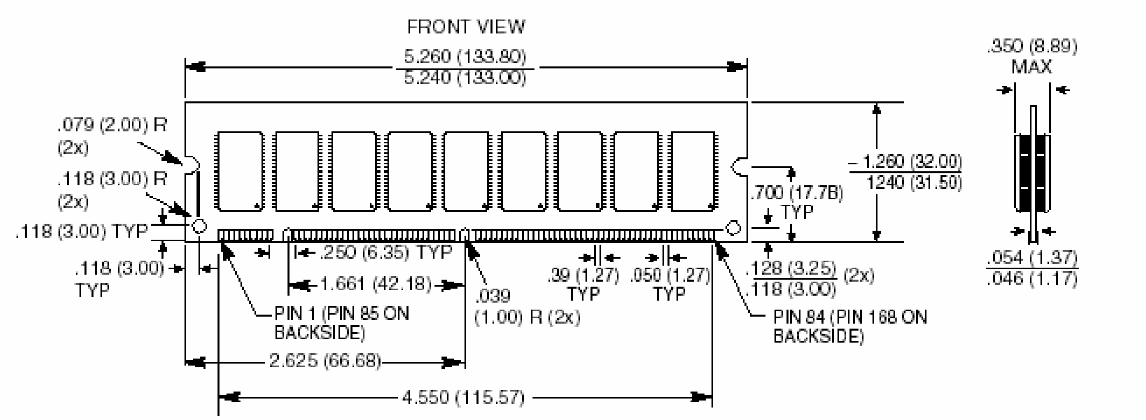
A typical 30-pin SIMM. The one shown here is 9-bit, although the dimensions would be the same for 8-bit.



A typical 72-pin SIMM. The version shown here is 36-bit, although the dimensions would be the same for 32-bit.

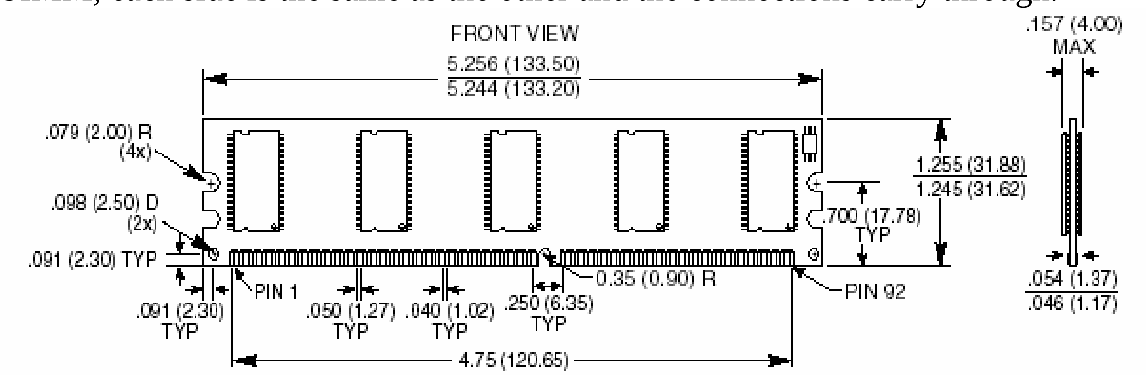
DIMM (Dual Inline Memory Module)

DIMMs are also available in two types:
Standard SDRAM
DDR SDRAM chips
Standard DIMMs have 168 pins, one notch on either side, and two notches along the contact area. DDR DIMMs, on the other hand, have 184 pins, two notches on each side, and only one notch along the contact area. All DIMMs are either 64-bits (nonparity) or 72-bits (parity or error correcting code [ECC]) wide (data paths). The main physical difference between SIMMs and DIMMs is that DIMMs have different signal pins on each side of the module. That is why they are called *dual* inline memory modules, and why with only an inch of additional length, they have many more pins than a SIMM.



A typical 168-pin SDRAM DIMM. The one shown here is 72-bit, although the dimensions would be the same for 64-bit

The pins are numbered from left to right and are connected through to both sides of the module on the SIMMs. The pins on the DIMM are different on each side, but on a SIMM, each side is the same as the other and the connections carry through.

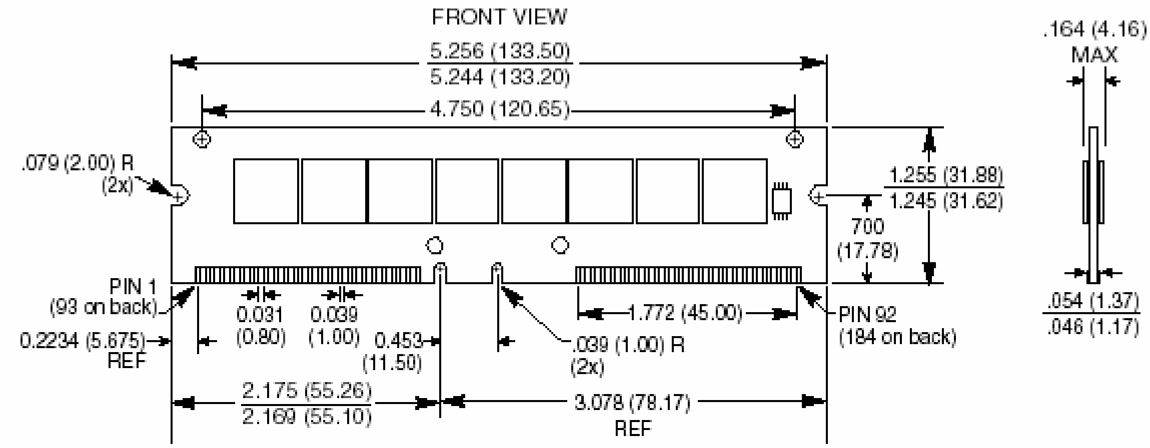


A typical 184-pin DDR DIMM. The one shown here is 72-bit, although the dimensions would be the same for 64-bit.

RIMM (RAMBUS INLINE MEMORY MODULE)

RIMM is a type of memory module made using RDRAM chips (See **Error! Reference source not found.**

RIMMs are also double-sided. Only one type of RIMM is available, which has 184 pins, one notch on either side, and two notches centrally located in the contact area. All RIMMs are either 16-bits wide or 18-bits wide (data path) for non-ECC and ECC versions, respectively.



A typical 184-pin RIMM. The one shown here is 18-bit, although the dimensions would be the same for 16-bit.

Gold- and Tin-plated Contacts of RAM

You must use the Gold-plated contact ram in the Gold-plated ram sockets on the motherboard and the Tin-plated contact ram in the Tin-plated ram socket on the ram. If you don’t follow the above rule, the ram would definitely work, but it might fail after a certain period normally one year. This is caused by a special type of corrosion called fretting. The fretting results from two no-similar metals coming together in contact under pressure. When you install the ram in the ram socket, the do fit in under considerable pressure.

If a ram failure is diagnosed, you should inspect the ram and corresponding socket contacts materials. If you observe layer of oxidized metal due to fretting, you may clean the contacts and bring the ram back to operational condition.

Parity and ECC

The ram is built of such components that it has to fail one day. The ram failures are of two types: hard fails and soft errors. The hard fails are easily understood. The chip fails because of some flaw , physical damage or some other event. This type of problem can only be fixed by replacing the failed component. Hard error rates are known as *HERs*. The other type of failure is the *soft error*, which is a nonpermanent failure that might never recur or could occur only at infrequent intervals. Soft fails can be removed by powering the system off and back on. Soft error rates are known as *SERs*.
The major causes of the soft errors are given bellow:

Power glitches or noise on the line. This can be caused by a defective power supply in the system or by defective power at the outlet.

- Incorrect type or speed rating. The memory must be the correct type for the chipset and match the system access speed.
- RF (radio frequency) interference. Caused by radio transmitters in close proximity to the system, which can generate electrical signals in system wiring and circuits.
- Static discharges. Causes momentary power spikes, which alter data.
- Timing glitches. Data doesn’t arrive at the proper place at the proper time, causing errors. Often caused by improper settings in the BIOS Setup, by memory that is rated slower than the system requires, or by over-clocked processors and other system components.

RAM Fault Tolerance

Three basic levels and techniques are used for fault tolerance in modern PCs:

- Non-parity
- Parity
- ECC

Non-parity

Non-parity systems have no fault tolerance at all. The only reason they are used is because they have the lowest inherent cost. No additional memory is necessary, as is the case with parity or ECC techniques. Because a parity-type data byte has 9 bits versus 8 for non-parity, memory cost is 12.5% higher.

Parity

Parity is a method of error checking in which an extra bit is sent to the receiving device to indicate whether an even or odd number of binary 1 bits was transmitted. The receiving unit compares the received information with this bit and can obtain a reasonable judgment about the validity of the character.

Nine memory chips in a bank handles 1 bit of data: 8 bits per character plus 1 extra bit called the parity bit. The parity bit enables memory-control circuitry to keep a check on the other 8 bits. This is a built-in cross-check for the reliability of each byte in the system. If the circuitry detects an error, the computer stops and displays a message informing you of the malfunction. If you are running a GUI operating system, such as Windows or OS/2, a parity error generally shows itself as a locked system. When you reboot, the BIOS should detect the error and display the appropriate error message. SIMMs and DIMMs are available both with and without parity bits. Originally, all PC systems used parity-checked memory to ensure accuracy.

ECC (Error Correcting Code)

ECC is much better than simple parity-error detection. Instead of just detecting an error, ECC allows a single bit error to be corrected, which means the system can continue without break and without corrupting data. ECC, as implemented in most PCs, can only detect, not correct, double-bit errors. Because studies have indicated that approximately 98% of memory errors are the single-bit variety, the most commonly used type of ECC is one in which the attendant memory controller detects and corrects single-bit errors in an accessed data word (double-bit errors can be detected, but not corrected).

Installing RAM Upgrades

Adding memory is a very useful and least expensive upgrade to a system. The working capacities of Windows 9x, Windows NT, Windows 2000, and OS/2 are increased by adding memory to the system. You can add the memory to the system in one of the following two ways:

- Adding memory in vacant slots on your motherboard
- Replacing your current motherboard’s memory with higher-capacity memory

Memory Banks

The memory chips are installed in memory banks. A *memory bank* is a collection of memory chips that make up a complete row of memory. Your processor reads each row of memory in one pass. A memory bank does not work unless the entire row is filled with memory chips.

Replacing Modules with Higher-Capacity Versions

If all the memory module slots on your motherboard are occupied, your best option is to remove an existing bank of memory and replace it with higher-capacity modules. For example, if you have two 64MB modules giving a total of 128MB, you could remove one of the 64MB modules and replace it with a 128MB unit, in which case you’d then have a total of 192MB of RAM. However, just because higher-capacity modules are available that are the correct pin count to plug into your motherboard, don’t automatically assume the higher-capacity memory will work. Your system’s chipset and BIOS set limits on the capacity of the memory you can use. Check your system or motherboard documentation to see which size modules work with it before purchasing the new RAM. You should make sure you have the latest BIOS when installing new memory.

Installing Memory

When you install or remove memory, you are most likely to encounter the following problems:

- ☛ Electrostatic discharge
- ☛ Improperly seated modules
- ☛ Incorrect memory configuration settings in the BIOS Setup

To prevent electrostatic discharge (ESD) when you install sensitive memory chips or boards, do not wear synthetic-fiber clothing or leather-soled shoes. Remove any static charge you are carrying by touching the system chassis before you begin, or better yet, wear a good commercial grounding strap on your wrist. A grounding strap consists of a conductive wristband grounded at the other end by a wire clipped to the system chassis. Leave the system unit plugged in—but turned off—to keep it grounded.

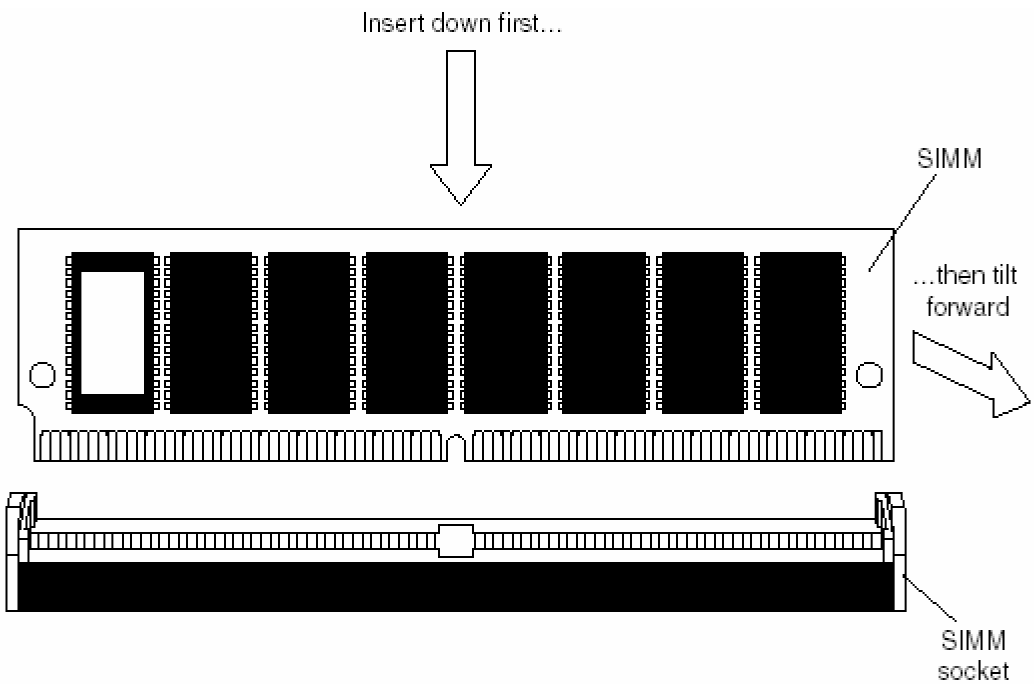
Warning
Be sure to use a properly designed commercial grounding strap; do not make one yourself. Commercial units have a one mega ohm resistor that serves as protection if you accidentally touch live power. The resistor ensures that you do not become the path of least resistance to the ground and therefore have an electrical shock. An improperly designed strap can cause the power to conduct through you to the ground, possibly killing you.

Each memory chip or module must be installed in a specific orientation. Observe the geometrical configuration and do not force anything. As you insert the module, be sure the notches align with the appropriate tangs on the socket. I would hope it goes without saying, but before installing memory, be sure the system power is off. You should make sure the system is unplugged just to be sure. If the system power was on when you installed the memory you could damage the memory and possibly the motherboard as well. You remove SIMMs and DIMMs by releasing the locking tabs and either pulling or rolling them out of their sockets. The installation is exactly the opposite.
After adding the memory and putting the system back together, you might have to run the BIOS Setup and resave with the new amount of memory being reported. Most new systems automatically detect the new amount of memory and reconfigure the BIOS Setup settings for you. Most new systems also don’t require setting any jumpers or switches on the motherboard to configure them for your new memory.

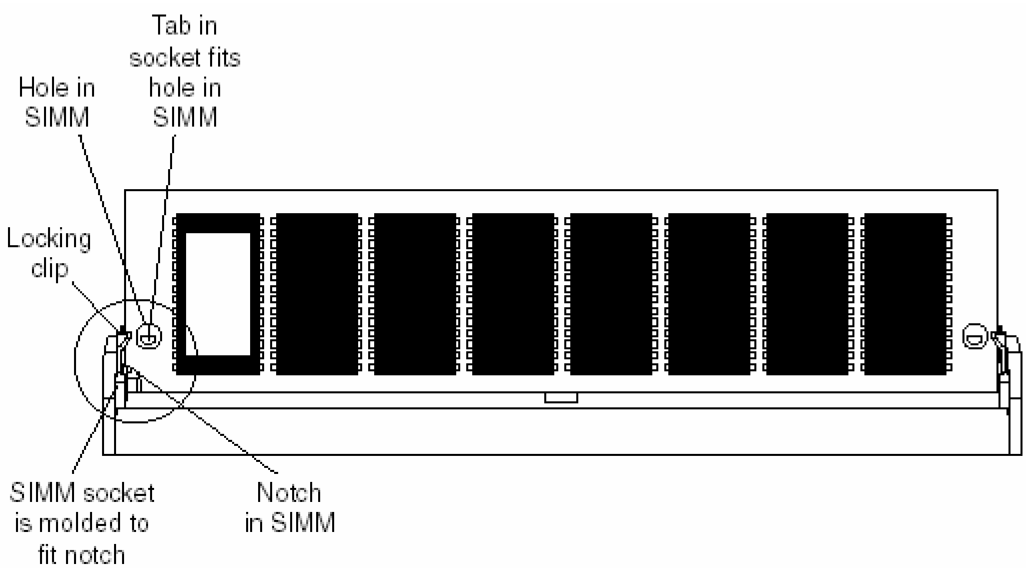
Installing SIMMs

SIMM memory is oriented by a notch on one side of the module that is not present on the other side, as shown in Figure 8.6. The socket has a protrusion that must fit into this

notched area on one side of the module. This protrusion makes installing a SIMM backward impossible unless you break the connector or the module. The picture below details the notch and the locking clip.



The notch on this SIMM is shown on the left side. Insert the SIMM at a 45° angle and then tilt it forward until the locking clips snap into place.



This picture shows the SIMM inserted in the socket with the notch aligned, the locking clip locked, and the hole in the SIMM aligned with the tab in the socket.

Installing DIMMs and RIMMs

DIMMs and RIMMs are keyed by notches along the bottom connector edge that are offset from the center so they can be inserted in only one direction, as shown in Figure 8.7. The DIMM/RIMM ejector tab locks into place in the notch on the side of the DIMM when it is fully inserted. Some DIMM sockets have ejector tabs on both ends. Take care not to force the module into the socket. If the module does not slip easily into the slot and then snap into place, there is a good chance the module is not oriented or aligned correctly. Forcing the module could break the module or the socket. If the retaining clips on the socket break, the memory isn't held firmly in place, and you likely will get random memory errors because the module can't make consistent electrical contact if it is loose.

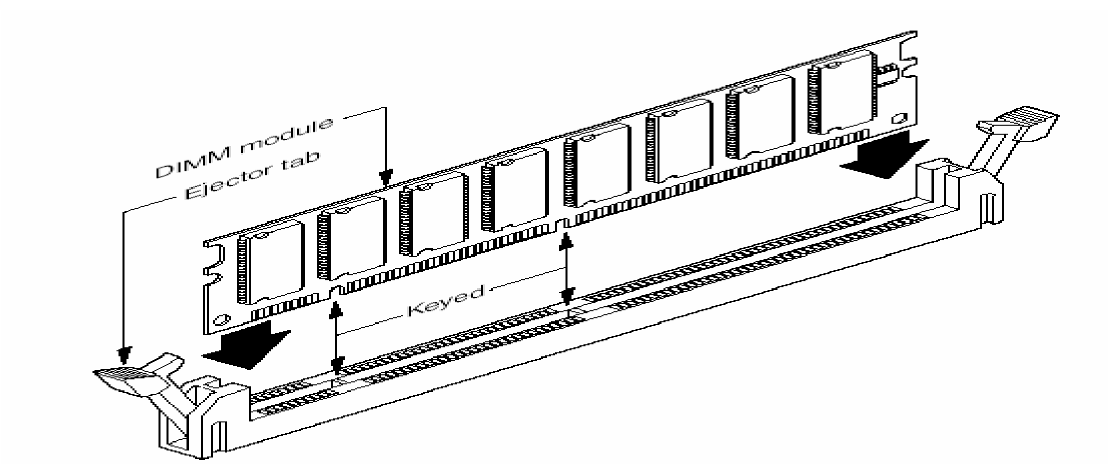
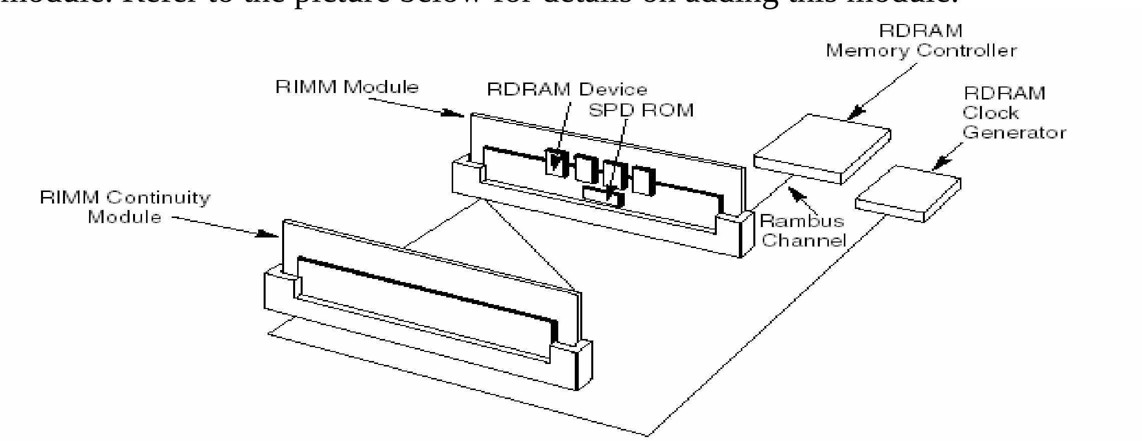


Figure 8. 1 DIMM keys match the protrusions in the DIMM sockets. DDR DIMM and RIMM keys are similar but not exactly the same.

Installing RIMMs

RIMMs install in exactly the same manner as DIMMs, although in different sockets. The ejector tabs on the sides are similar, and they go in at a straight 90° angle just like DIMMs. The key notches on the RIMM prevent backwards installation and aid alignment when inserting them. You'll also need to fill any empty RIMM sockets with a continuity module. Refer to the picture below for details on adding this module.



Typical RDRAM bus layout showing a RIMM and one continuity module

Troubleshooting Memory

To troubleshoot memory, you first need some memory-diagnostics testing programs. Every motherboard BIOS has a memory diagnostic in the POST that runs when you first turn on the system. When the POST runs, it not only tests memory, but also counts it. The count is compared to the amount counted the last time BIOS Setup was run; if it is different, an error message is issued. If any failure is detected, you see or hear a message. Audio messages (beeping) are used for critical or “fatal” errors that occur in areas important for the system’s operation. If the system can access enough memory to at least allow video to function, you see error messages instead of hearing beep codes.

If your system makes it through the POST with no memory error indications, there might not be a hardware memory problem, or the POST might not be able to detect the problem. Intermittent memory errors are often not detected during the POST, and other subtle hardware defects can be hard for the POST to catch. The POST is designed to run quickly, so the testing is not nearly as thorough as it could be.

After your operating system is running, memory errors can still occur, typically identified by error messages you might receive. These are the most common:

- Parity errors. Indicates that the parity-checking circuitry on the motherboard has detected a change in memory since the data was originally stored.
- General or global protection faults. A general-purpose error indicating that a program has been corrupted in memory, usually resulting in immediate termination of the application. This can also be caused by buggy or faulty programs.
- Fatal exception errors. Error codes returned by a program when an illegal instruction has been encountered, invalid data or code has been accessed, or the privilege level of an operation is invalid.
- Divide error. A general-purpose error indicating that a division by 0 was attempted or the result of an operation does not fit in the destination register.

If you are encountering these errors, they could be caused by defective or improperly configured memory, but they can also be caused by software bugs (especially drivers), bad power supplies, static discharges, close proximity radio transmitters, timing problems, and more.

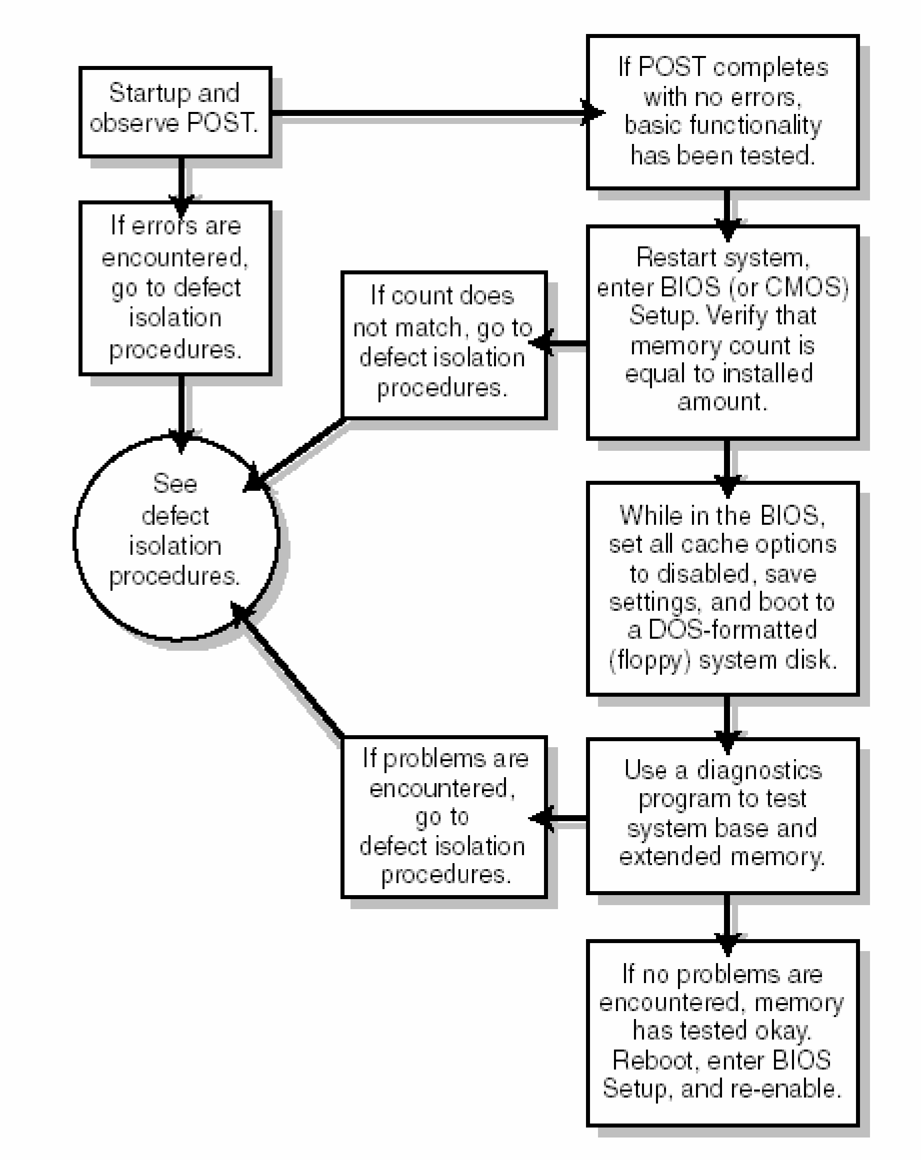
If you suspect the problems are caused by memory, there are ways to test the memory to determine whether that is the problem. Most of this testing involves running one or more memory test programs.

Warning
When you run the memory tests with the system caches enabled. This makes the memory testing invalid because most systems have what is called a *write-back cache*. This means that data written to main memory is first written to the cache. Because a memory test program first writes data and then immediately reads it back, the data is read back from the cache, not the main memory. It makes the memory test program run very quickly, but all you tested was the cache. The bottom line is that if you test memory with the cache enabled, you aren't really writing to the SIMM/DIMMs, but only to the cache. Before you run any memory test programs, be sure your cache is disabled. The system will run very slowly when you do this, and the memory test will take much longer to complete, but you will be testing your actual RAM, not the cache.

The following steps enable you to effectively test and troubleshoot your system RAM. Figure 8.10 provides a boiled-down procedure to help you step through the process quickly.

First, let's cover the memory testing and troubleshooting procedures.
Power up the system and observe the POST. If the POST completes with no errors, basic memory functionality has been tested. If errors are encountered, go to the defect isolation procedures.

- Restart the system, and enter your BIOS (or CMOS) Setup. In most systems, this is done by pressing the F2 key during the POST but before the boot process begins. Once in BIOS Setup, verify that the memory count is equal to the amount that has been installed. If the count does not match what has been installed, go to the defect isolation procedures.
- Find the BIOS Setup options for cache, and set all cache options to disabled. Save the settings and reboot to a DOS-formatted system disk (floppy) containing the diagnostics program of your choice. If your system came with a diagnostics disk, you can use that, or you can use one of the many commercial PC diagnostics programs on the market, such as PC-Technician by Windsor Technologies (which comes in self-booting form), the Norton Utilities by Symantec, or others.



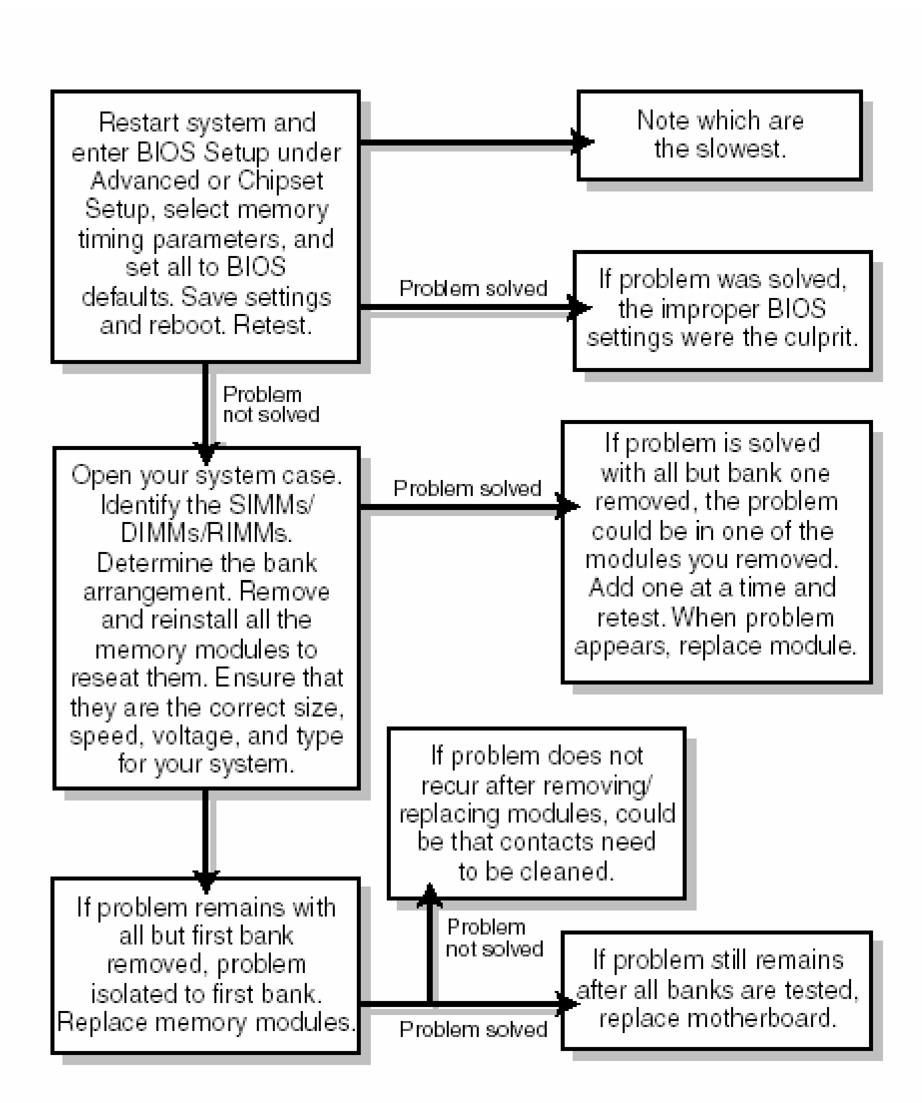
Testing and troubleshooting memory

- Follow the instructions that came with your diagnostic program to have it test the system base and extended memory. Most programs have a mode that enables them to loop the test—that is, to run it continuously, which is great for finding intermittent problems. If the program encounters a memory error, proceed to the defect isolation procedures.
- If no errors are encountered in the POST or in the more comprehensive memory diagnostic, your memory has tested okay in hardware. Be sure at this point to reboot the system, enter the BIOS Setup, and re-enable the cache. The system will run very slowly until the cache is turned back on.
- If you are having memory problems yet the memory still tests okay, you might have a problem undetectable by simple pass/fail testing, or your problems could be caused by

software or one of many other defects or problems in your system. You might want to bring the memory to a SIMM/DIMM tester for a more accurate analysis. Most PC repair shops have such a tester. I would also check the software (especially drivers, which might need updating), power supply, and system environment for problems such as static, radio transmitters, and so forth.

Memory Defect Isolation Procedures

The following procedure isolates which SIMM or DIMM in the system is causing the problem.



- Restart the system and enter the BIOS Setup. Under a menu usually called Advanced or Chipset Setup might be memory timing parameters. Select BIOS or Setup defaults,

which are usually the slowest settings. Save the settings, reboot, and retest using the testing and troubleshooting procedures listed earlier. If the problem has been solved, improper BIOS settings were the problem. If the problem remains, you likely do have defective memory, so continue to the next step.

- Open the system for physical access to the SIMM/DIMMs on the motherboard. Identify the bank arrangement in the system. For example, Pentium systems use 64-bit banks, which means two 72-pin SIMMs or one 168-pin DIMM per bank. Using the manual or the legend silkscreened on the motherboard, identify which SIMMs or DIMM correspond to which bank. Most Pentium boards that use SIMMs have four, six, or eight total SIMMs installed in two, three, or four banks. Most boards using DIMMs have two, three, or four total DIMMs, each of which represents a single bank.
- Remove all the memory except the first bank, and retest using the troubleshooting and testing procedures listed earlier. If the problem remains with all but the first bank removed, the problem has been isolated to the first bank, which must be replaced.
- Replace the memory in the first bank, preferably with known good spare modules, but you can also swap in others that you have removed and retest. If the problem *still* remains after testing all the memory banks (and finding them all to be working properly), it is likely the motherboard itself is bad (probably one of the memory sockets). Replace the motherboard and retest.
- At this point, the first (or previous) bank has tested good, so the problem must be in the remaining modules that have been temporarily removed. Install the next bank of memory and retest. If the problem resurfaces now, the memory in that bank is defective. Continue testing each bank until you find the defective module.
- Repeat the preceding step until all remaining banks of memory are installed and have been tested. If the problem has not resurfaced after removing and reinstalling all the memory, the problem was likely intermittent or caused by poor conduction on the memory contacts. Often simply removing and replacing memory can resolve problems because of the self-cleaning action between the SIMM/DIMM and the socket during removal and reinstallation.

Logical Memory Layout

The logical memory layout of the original PC was as follows:

Total addressable memory	10. 01 MB	11. (1024 KB)
The memory to be used by the applications	12. 1st 640 KB	13. (0 KB – 640 KB)
The memory to be used by the system (Video Ram), other adapters (for on-card ROM BIOS), and finally the motherboard ROM BIOS	14. Last 384 KB	15. (641 KB – 1024 KB)

All the PCs up to Pentium 4 use the same scheme when using a 16-bit Operating System occasionally. Since most people run 32-bit operating systems, such as Windows 9x, 2000, Linux, and so on, these operating systems automatically manage the use of RAM. The following sections are intended to give you an understanding of the PC hardware memory layout, which is consistent no matter which operating system you use. The only things that change are how your operating system uses these areas and how they are managed by the OS.

The following sections detail the various types of memory installed on a modern PC. The kinds of memory covered in the following sections include:

- Conventional (base) memory
- Upper memory area (UMA)
- Extended memory

If the processor is running in real mode, only the first megabyte is accessible. If the processor is in protected mode, the full memory is accessible.

Conventional (Base) Memory

The first megabyte or first 640KB of system memory accessible by an Intel processor in real mode is called the conventional or base memory. The barrier of base memory basically affects 16-bit software, such as DOS and Windows 3.1, and is much less of a factor with 32-bit software and operating systems, such as Windows 9x, and NT/2000 etc.

Upper Memory Area

The term *Upper Memory Area (UMA)* describes the reserved 384KB at the top of the first megabyte of system memory on a PC/XT and the first megabyte on an AT-type system.

Extended Memory

Extended memory is basically all memory past the first megabyte, which can be accessed only while the processor is in protected mode.

Chapter 8

Basic Disk Drives

OBJECTIVES

Floppy Disk Drives

Hard Disk Drives

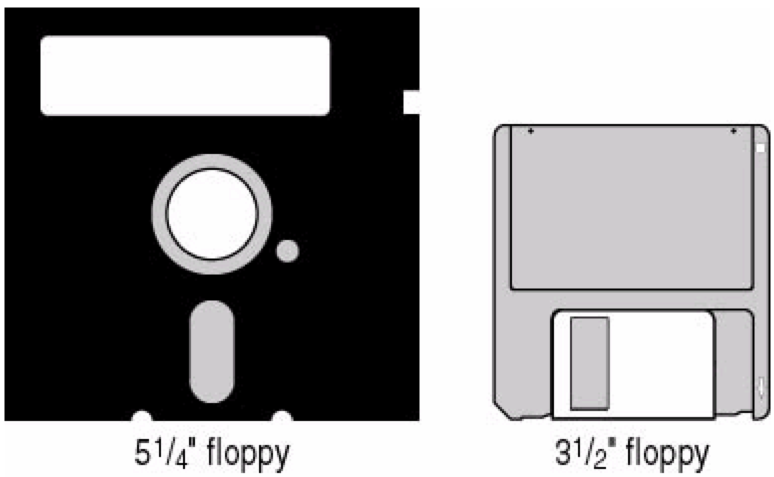
Other PC Peripheral

Advanced Hard Disk Drives

Floppy Disk Drives

The Basics of Floppy Disk Drives

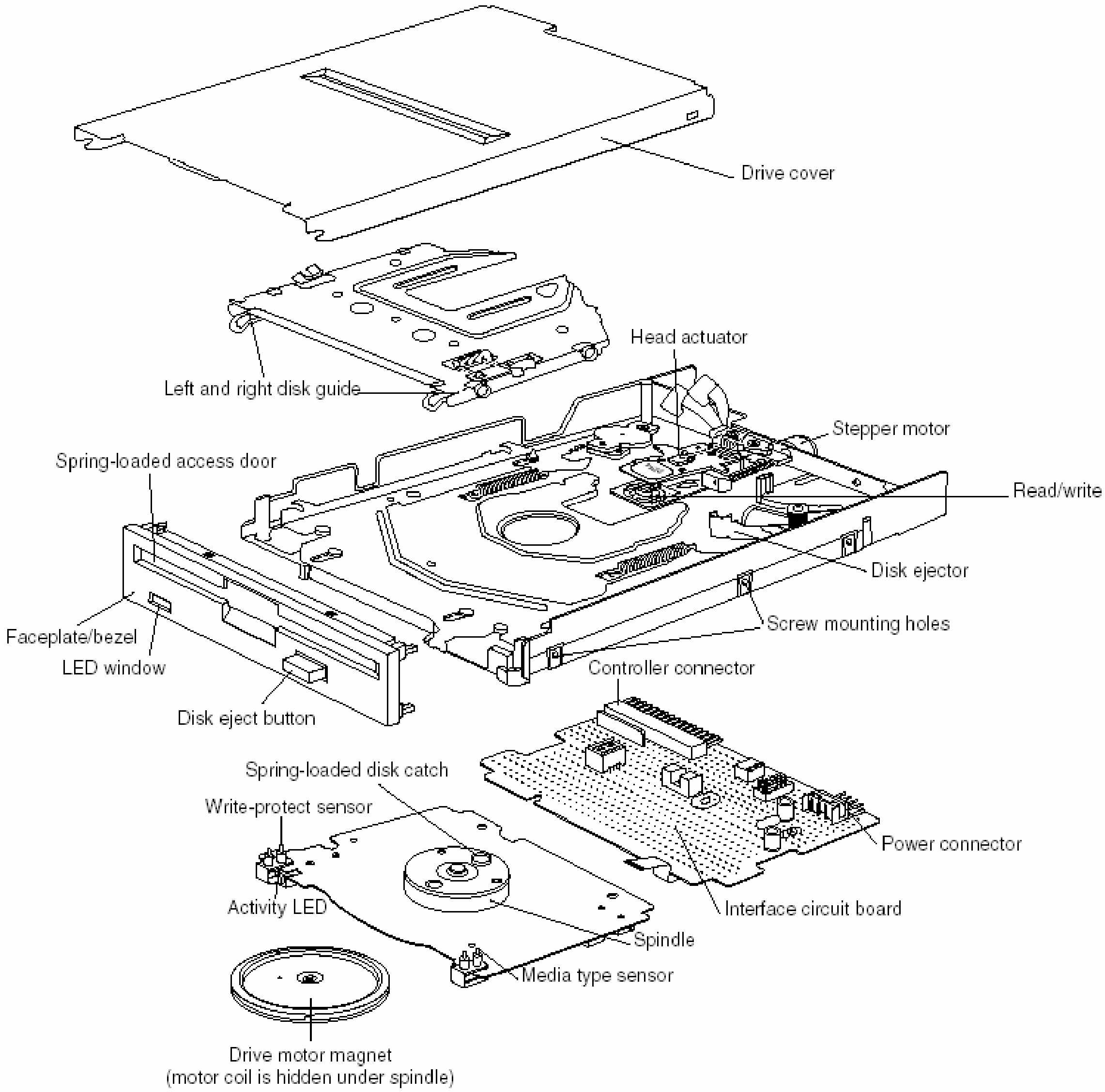
Today's floppy disks are made of flexible plastic and coated with a magnetic material. To protect the disk from dust and physical damage, it is packaged in a plastic or coated paper case. The main reason for the popularity of floppy disk drives and disks is that they provide inexpensive read/write (R/W) removable media. The data stored on a floppy disk can be moved from one computer to another, provided both have the same type of drive. In general, it is a good idea to protect your data by always keeping two copies of any data file that you create (the original and a backup), and the floppy disk is an excellent medium for backing up, storing, or distributing copies of relatively small files, such as word-processing documents.



Floppy disks

Drive Components

All floppy disk drives, regardless of type, consist of several basic common components. To properly install and service a disk drive, you must be able to identify these components and understand their functions (see Figure 10.1).



A typical floppy disk drive.

Floppy Drive Controller

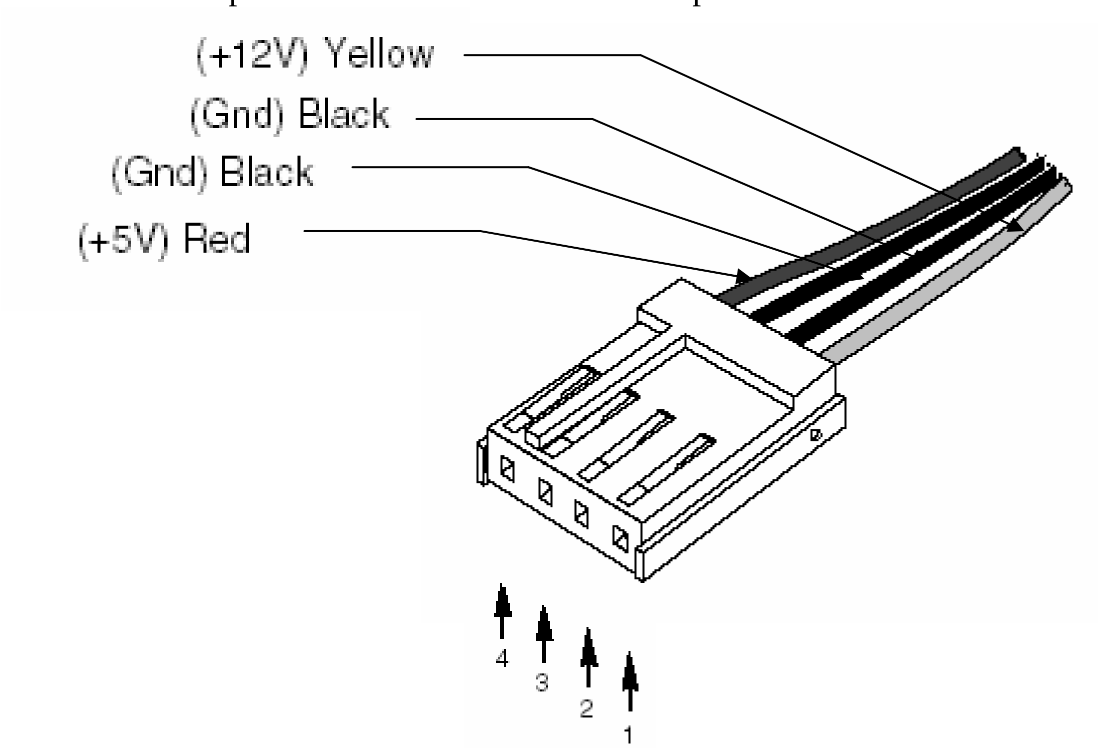
Today's PCs have the floppy controller integrated into the motherboard, usually in the form of a Super I/O chip that also includes the serial and parallel interfaces, among other things. Even though the floppy controller can be found in the Super I/O chip on the motherboard, it is still interfaced to the system via the ISA bus and functions exactly as if it were a card installed in an ISA slot. These built-in controllers are normally configured via the system BIOS Setup routines and can be disabled if an actual floppy controller card is going to be installed.

Whether it is built in or not, each primary floppy controller uses a standard set of system resources:

- IRQ 6 (Interrupt Request)
- DMA 2 (Direct Memory Address)
- I/O ports 3F0–3F5, 3F7 (input/output)

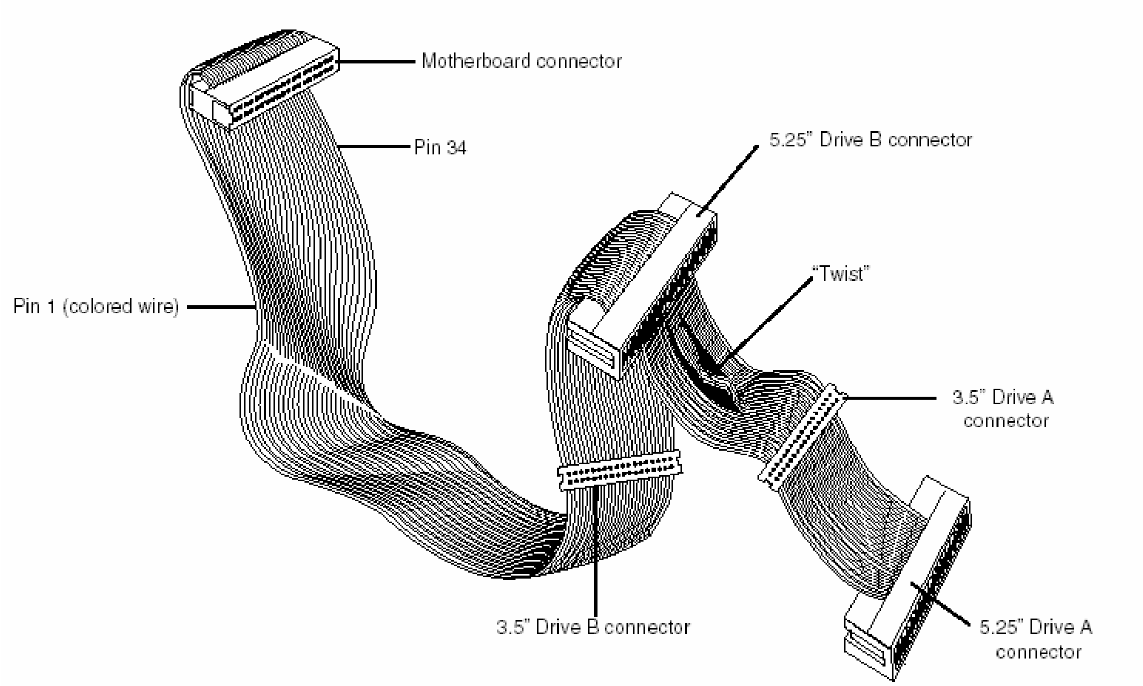
Connectors

Floppy disk drives have two connectors. One of them is for power to run the drive and the other to carry the control and data signals to and from the drive. These connectors are standardized in the computer industry. PCs use 3 1/2-inch drives with a 34-pin signal connector and a power connector as shown in the picture below



3.5” floppy disk drive female power supply cable connector

The Floppy Disk Drive Cable



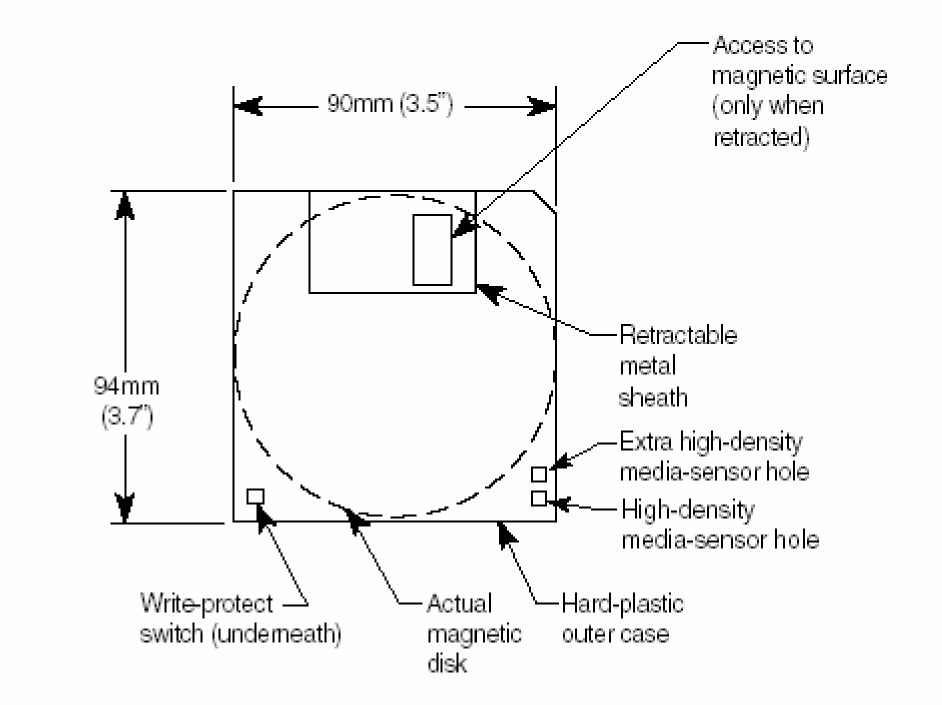
Standard five-connector floppy interface cable.

Floppy Disk Drive Operation

Most PCs sold today are equipped with a 3 1/2-inch 1.44MB floppy disk drive.
The physical operation of a disk drive is fairly simple to describe. The disk rotates in the drive at either 300rpm. With the disk spinning, the heads can move in and out approximately one inch and write 80 tracks. The tracks are written on both sides of the disk and are therefore sometimes called *cylinders*.

Floppy Disk Drive Construction

The 3 1/2-inch disks use a semi-rigid plastic case, which helps stabilize the magnetic medium inside. A metal shutter protects the media-access hole. The drive manipulates the shutter, leaving it closed whenever the disk is not in a drive. The medium is then completely insulated from the environment and from your fingers. The shutter makes the disk jacket unnecessary.



Construction of a 3 1/2-inch floppy disk

On the lower-left part of the disk is a hole with a plastic slider—the write-protect/enable hole. When the slider is positioned so the hole is visible, the disk is *write-protected*—the drive is prevented from recording on the disk. When the slider is positioned to cover the hole, writing is enabled, and you can save data to the disk. The magnetic medium in 3 1/2-inch disks is constructed of a plastic base (usually Mylar) coated with a magnetic compound.

Floppy disks can be damaged or destroyed easily by the following:

- Touching the recording surface with your fingers
- Writing on a disk label
- Bending the disk
- Spilling coffee or other substances on the disk
- Overheating a disk
- Exposing a disk to magnetic fields

Drive-Installation Procedure

Floppy disk drives are installed into the same half-height bays as hard disk drives; the physical mounting of the drive in the computer case is the same for both units. See Unit 9 Hard Disk Drive.

Troubleshooting Floppy Drives

We will discuss some of the most common problems floppy drives can have and how to troubleshoot them.

Problem

Dead drive—the drive does not spin and the LED never comes on.

Cause/Solution

The drive or controller is not properly configured in the BIOS setup. Check the BIOS setup for proper drive type and make sure the controller is enabled if it's built in to the motherboard.
Other causes for a dead drive include

- *Bad power supply or power cable.* Measure the power at the cable with a voltmeter, and ensure that 12v and 5v are available to the drive.
- *Bad data cable.* Replace the cable and retest.
- *Defective drive.* Replace the drive and retest.
- *Defective controller.* Replace the controller and retest. If the controller is built in to the motherboard, disable it via the BIOS setup, install a card-based controller, and retest, or replace the entire motherboard and retest.

Problem

Drive LED remains on continuously.

Cause/Solution

Data cable is on backward at either the drive or controller connection. Reinstall the cable properly and retest.
Also, the data cable could be offset on the connector by one or more pins. Reinstall the cable properly and retest. If this does not solve the problem, it could be a defective cable. Replace the cable and retest.

Problem

Ghost directories—You have exchanged disks in the drive, but the system still believes the previous disk is inserted, and even shows directories of the previous disk.

Cause/Solution

Several causes and solutions are possible:

- *Defective cable (pin 34)*. Replace the cable and retest.
- *Improper drive configuration*. If the drive is an older model and has a DC jumper, it must be set to enabled. You also should ensure the BIOS Setup has the proper drive type entered.
- *Defective drive*. Replace the drive and retest.

Common Floppy Drive Error Messages—Causes and Solutions

Error Message

Invalid Media or Track Zero Bad, Disk Unusable

Cause/Solution

You are formatting the disk and the disk media type does not match the format parameters. Make sure you are using the correct type of disk for your drive and formatting the disk to its correct capacity.
Other causes and solutions are

- *Defective or damaged disk.* Replace the disk and retest.
- *Dirty heads.* Clean the drive heads and retest.

Error Message

CRC Error or Disk Error 23

Cause/Solution

The data read from the disk does not match the data that was originally written. (CRC stands for cyclic redundancy check.) Replace the disk and retest, or possibly, clean the drive heads. Use the Norton Utilities for possible data recovery from the disk.

Error Message

General Failure Reading Drive A, Abort, Retry, Fail or Disk Error 31

Cause/Solution

The disk is not formatted or has been formatted for a different operating system (Macintosh, for example). Reformat the disk and retest. Another cause could be that damaged areas exist on the disk medium. Replace the disk and retest. Use the Norton Utilities for possible data recovery from the disk.

Error Message

Access Denied

Cause/Solution

You are trying to write to a write-protected disk or file. Move the write-protect switch to allow writing on the disk, or remove the read-only file attribute from the files. File attributes can be changed with the ATTRIB command.

Error Message

Insufficient Disk Space or Disk Full

Cause/Solution

The disk is filled, or the root directory is filled. Check to see whether sufficient free space is available on the disk for your intended operation.

Error Message

Bytes in Bad Sectors

Cause/Solution

This is displayed after formatting or running the CHKDSK command and shows that the disk has several allocation units that are marked bad. This is not a problem in itself because when the allocation units are marked in this fashion, the operating system will not use them for file storage. Even so, I normally don't recommend using disks with bad sectors.

Error Message

Disk Type or Drive Type Incompatible or Bad

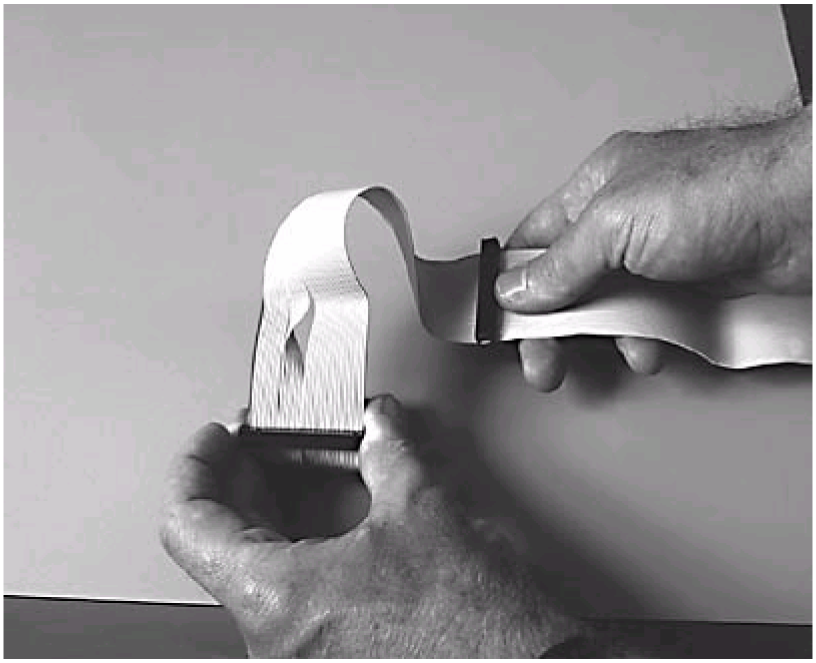
Cause/Solution

You are attempting to DISKCOPY between two incompatible drive or disk types. Disks can be copied only between drives using the same disk density and size.

The following table describes various floppy disks and their capacities.

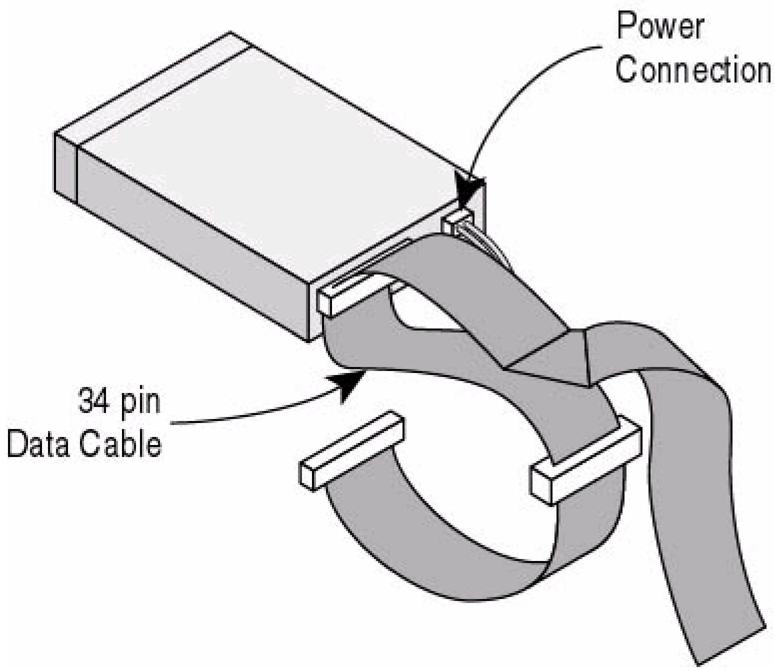
Disk Size	Capacity	Description
5.25-inch	160 KB	Single-sided single-density—the first model.
5.25-inch	360 KB	Double-sided single-density.
5.25-inch	720 KB	Double-sided double-density.
5.25-inch	1.2 MB	Double-sided high-density.
3.5-inch	720 KB	Double-sided double-density.
3.5-inch	1.44 MB	Double-sided high-density—today's standard.
3.5-inch	2.88 MB	Double-sided quad-density. This format has never really gained in market share and is not common on today's PCs.

All floppy disk drives are connected to the motherboard's external data bus by a 34-conductor ribbon cable, shown in the picture below. This cable has a seven-wire twist in lines 10 through 16. This ensures that when two floppy disk drives are attached, the drive-select and motor-enable signals on those wires can be inverted to "select" which drive becomes the active target. The other wires carry data and ground signals. The connector end of the cable with the twist *always* goes toward the drives.



Floppy disk drive cable with a twist

If you install a new drive and notice that the indicator light comes on and stays on, the cable is most likely backward. The power connection for a floppy disk drive, shown in the picture below, is either the large Molex type connector on the 5.25-inch drive or the smaller mini connector, on the 3.5-inch drive. A large number of power supplies, designed primarily for tower systems, provide a special connector exclusively for 3.5-inch floppy disk drives. This connector is a two-strand connection that provides a 5-volt power connection to the 3.5-inch drive.



Floppy disk drive cable connections

After you physically install a floppy disk drive, you need only set the proper CMOS settings for the type and position (first or second) and the installation will be complete. In CMOS setup, select the drive (A or B) and enter the correct capacity. (CMOS stands for complementary metal-oxide semiconductor.)

Many older CMOS chips won't have settings for 1.44 MB or 2.88 MB 3.5-inch floppy disk drives because they were developed before these drives were introduced. Also, the 5.25-inch drives are virtually obsolete, and the CMOS of the future might not have settings for them. Several third-party utilities will allow the CMOS to accept the necessary values to support these drives.

Keeping a Floppy Disk Drive Running

Although floppy disk drives are usually rugged and dependable, they do take a lot of abuse and sometimes they fail. Some failures are simply caused by improper use, some by overuse combined with a lack of cleaning, and sometimes the mechanism just stops working.

Floppy disk drives fail more than any other part of a computer system. They are highly susceptible to failure because their internal components are directly exposed to the outside world. Often, there is only a small door or slot that separates the R/W heads from dust, grime, and cigarette smoke. Floppy disk drives are often the victims of inverted disks, paper clips, and other foreign objects that can cause mechanical damage.

Floppy disk drives are inexpensive and easy to replace. The only preventive maintenance required is to keep the floppy disk drive *clean*! Excellent cleaning kits are available in most computer and discount stores. To achieve the best performance from a floppy disk drive in a high-use or industrial environment, schedule monthly cleaning.

Always an Exception

One unusual floppy disk drive solution that appeared as the 3.5-inch models gained dominance was the hybrid 3.5/5.25 drive. This married the slots for both formats in a single housing. They install just like a single drive.

Errors Caused by the Floppy Disk

If a floppy disk drive doesn't work, the first thing you suspect should be the floppy disk.

To check a floppy disk, follow this procedure:

1. First, make sure the disk is not write protected. The hole on the right top corner of a 3.5-inch disk (viewed from the front) should be closed. On a 5.25-inch disk, the notch on one side should be visible (not covered).
2. Try another disk.
3. Try a new (formatted) disk.
4. Try someone else's disk—one that is known to work on another computer (first make sure there is no critical data on the disk).
5. If two or more disks are unreadable, the drive is suspect; try going to MS-DOS and reading a directory using the DIR command.

CAUTION

Never test a drive by using a disk that contains important data! If the drive is bad, it may destroy any disks placed into it.

Detecting Data Errors on a Disk

If you can read data from one disk, but not another, the problem is the floppy disk. *Throw it away*. Data errors on floppy disks generally result in an error message that ends with the words "Abort, Retry, Fail." The process for repairing floppy disks is identical to the process for repairing hard disk drives drives, should there be data on the disk that must be recovered.

Check the CMOS Setting

Occasionally, the CMOS settings for floppy disks cause problems with drive operations. Any of the following errors indicates a possible CMOS setup problem:

- General failure reading drive A: (or B:)
- Not ready error reading drive A: (or B:)
- Insert disk for drive A: (or B:) and press any key when ready

BIOS makers often use the 3.5-inch high-density disk drive as the default CMOS setting for the A drive. With this BIOS, failure of the CMOS battery, or even accidental erasure of the CMOS, will still allow most floppy disks to work. Always double-check the CMOS if you are experiencing a recurrent floppy disk drive failure. It is quick, easy, and might save you time.

Check/Change the Floppy Disk Drive Cable

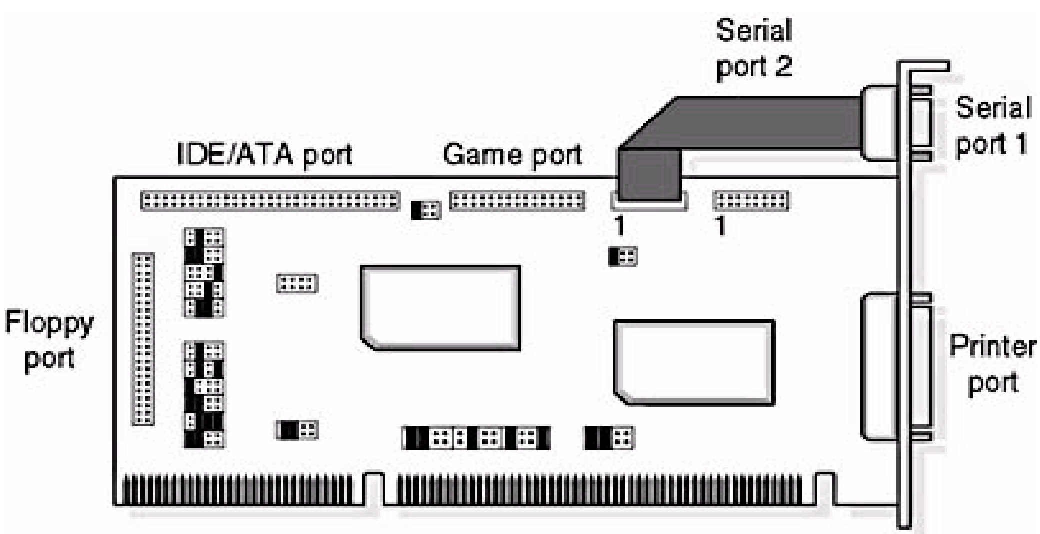
Cables wear out, work themselves loose, and are sometimes improperly installed. Check out both the data cable and the power jack as causes of the errant floppy disk drive before moving on to the controller.

Change the Floppy Disk Drive Controller

Floppy disk drive controllers are durable and highly resistant to failure and therefore should be the *last* components to blame. Left alone, they generally cause no problems. However, if they have recently been handled, such as during a move or repairs to the computer, they can be suspect. They are extremely sensitive to shock and static discharge.

In the event of a loose data cable or power plug, the POST (power-on self test) will return "FDD Controller Failure" or "Drive Not Ready" errors. Verify all the connections and try again. If the connections are good, try removing and reseating the controller (beware of electrostatic discharge). If the same errors continue, replace the controller. Floppy disk drives and controllers are inexpensive.

When replacing a floppy disk drive controller, keep in mind that most of these controllers on pre-Pentium machines are bundled as part of a combination I/O card. These cards include some (often all) of the following: hard disk drive controllers, serial ports, parallel ports, and joystick ports. If the new card contains any duplicate ports (they already exist elsewhere on the computer), a potential for conflict will exist.



I/O card with floppy disk controller

TIP

Be sure to disable all other devices on the card that are already installed on the computer when installing I/O cards. If not disabled, the I/O card can cause conflicts. If you have a new card with improved devices, disable or remove the older item.

Replace the Floppy Disk Drive

When replacing floppy disk drives, be sure to throw away the old drive. Floppy disk drives are inexpensive compared to other components in the computer. Consider purchasing them in quantity and saving money. It is not a bad idea to have a spare floppy disk drive and I/O card available for testing purposes.

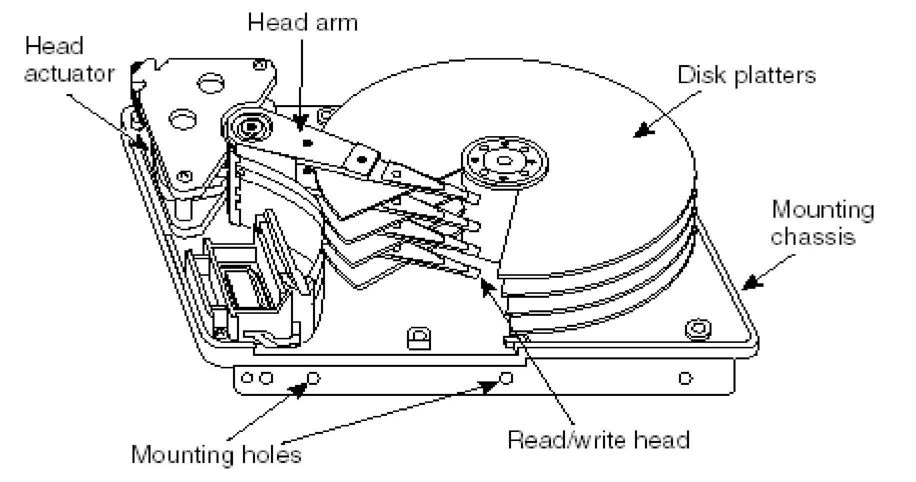
NOTE

As a rule, floppy disk drives fail more than any other part of a computer system except the floppy disk itself.

Hard Disk Drive

Definition of a Hard Disk

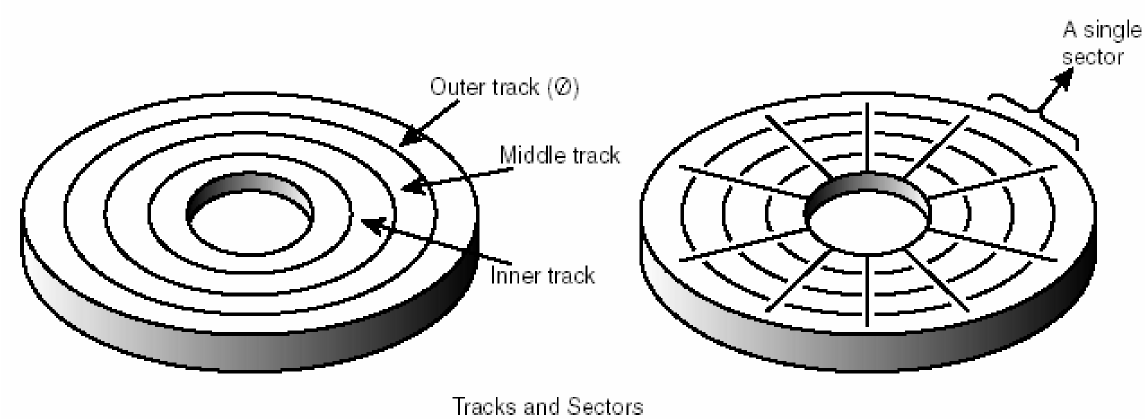
Hard Disk Drive may be defined as PC component that is used for the permanent storage of data. The data stored in the hard disk is not lost by power failure. Hard drives have rigid and round disk which are called platters. These platters are made of aluminum or glass. This is because of these rigid platters that the term hard disk is used. These platters cannot be removed. This is why they are sometimes called *fixed* disk drives.



Hard disk heads and platters.

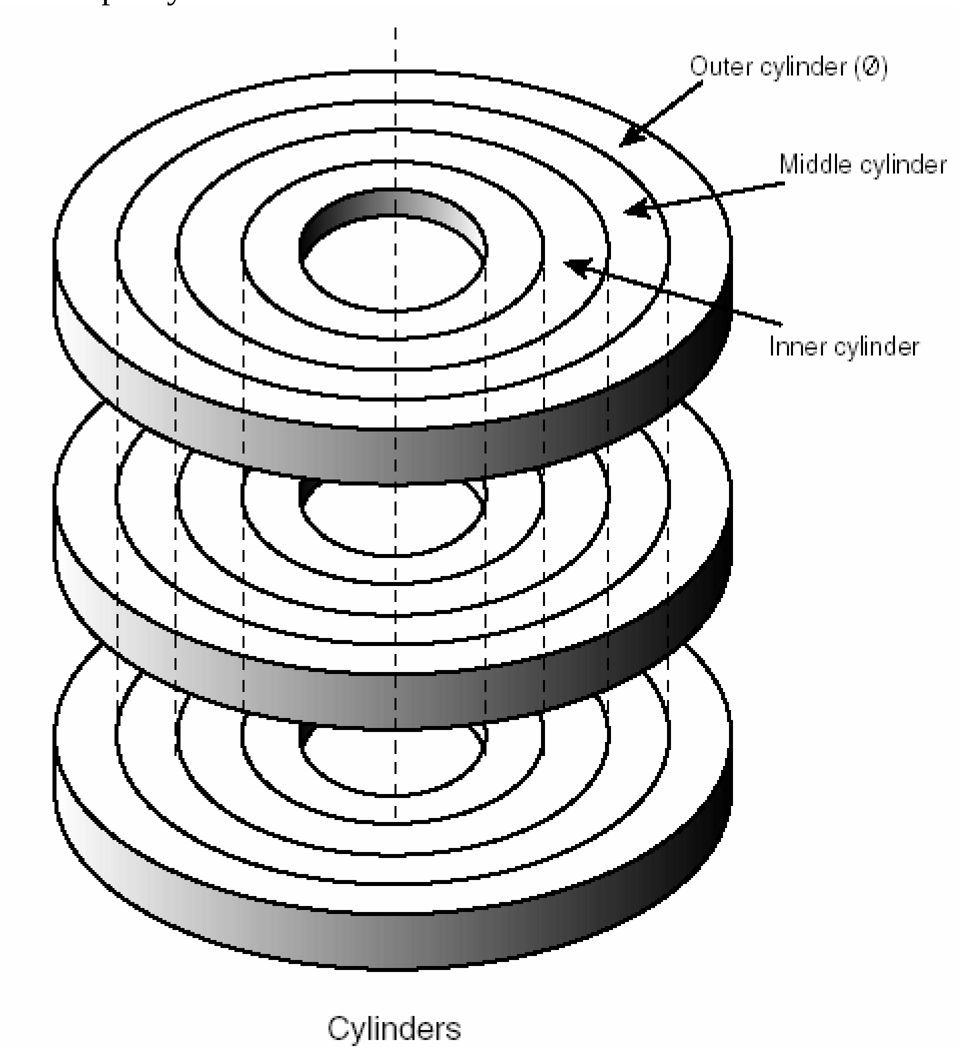
Hard Disk Drive Construction

The two main components of the hard disk drive are the revolving disks and the heads which read/write the data on the disks. The data is read and written on tracks as shown bellow. The tracks are further divided into *sectors*, which normally store 512 bytes each.



The tracks and sectors on a disk.

The platters are stacked on top of each other and rotate together. Each platter has two sides on which the drive stores data. Most drives have two or three platters, resulting in four or six sides. The identically aligned tracks on each side of every platter together make up a cylinder.



Hard disk cylinders

All the heads are mounted on a common carrier device or rack. The heads move radially across the disk together. They cannot move independently because they are mounted on the same carrier or rack, called an *actuator*. Originally, most hard disks spun at 3,600rpm—approximately 10 times faster than a floppy disk drive. Most of the modern drives spin the platters at 4,200; 5,400; 7,200; 10,000; or 15,000rpm. High rpm disks are normally faster than lower rpm drives. The heads in most hard disk drives do not (and should not!) touch the platters during normal operation.
When the heads are powered off, however, in most drives they land on the platters as they stop spinning. While the drive is running, a very thin cushion of air keeps each head suspended a short distance above or below the platter. This air gap is in micro inches.

Tracks

A *track* is a single ring of data on one side of a disk.

Sectors

Tracks are divided into several numbered divisions known as sectors. These sectors represent arc-shaped pieces of the track. Different type of disk drives split their disk tracks into different numbers of sectors, depending on the density of the tracks. Hard disks can have 900 or more sectors per track physically. Each sector contains 512 bytes of data.

Disk Formatting

You have to perform following three steps to make the drive ready for data storage:

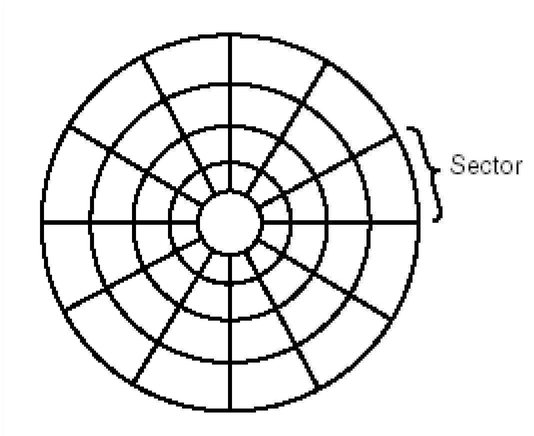
- Low-level formatting (LLF)
- Partitioning
- High-level formatting (HLF)

Low Level Formatting

During a low-level format, the formatting program divides the disk's tracks into a specific number of sectors, creating the intersector and intertrack gaps and recording the sector header and trailer information. The program also fills each sector's data area with a dummy byte value or a pattern of test values. The number of sectors per track depends on the drive and the controller interface.

Standard Recording

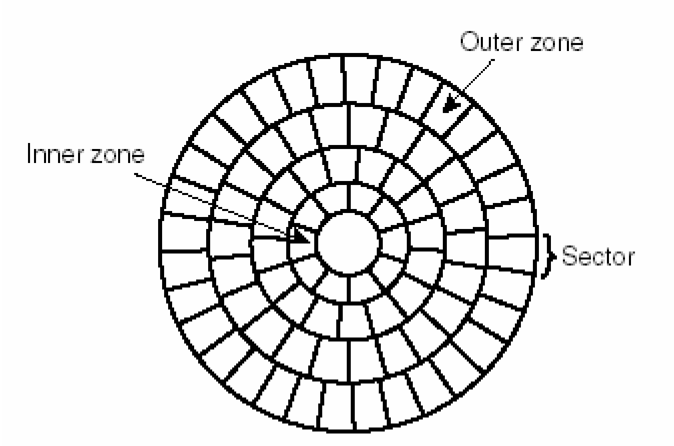
A standard recording wastes capacity on the outer tracks, because it is longer and yet holds the same amount of data (more loosely spaced) as the inner tracks. One way to increase the capacity of a hard drive during the low-level format is to create more sectors on the disks' outer cylinders than on the inner ones. Because they have a larger circumference, the outer cylinders can hold more data. Drives without zoned-bit recording store the same amount of data on every cylinder, even though the tracks of the outer cylinders might be twice as long as those of the inner cylinders. The result is wasted storage capacity because the disk medium must be capable of storing data reliably at the same density as on the inner cylinders. When the number of sectors per track is fixed, as in older controllers, the drive capacity is limited by the density of the innermost (shortest) track.



Standard recording, where the same number of sectors comprise every track

Zoned-Bit Recording

Virtually all ATA and SCSI drives use a technique called *zoned-bit recording*, which writes a variable number of sectors per track. Without zoned-bit recording, the number of sectors, and therefore bits, on each track is a constant.



Zoned-bit recording, where the number of sectors per track increases within each zone, moving out from the center

Low-level formats of ATA and SCSI hard disk drives are performed by the manufacturer and should almost never be performed by the end user. The only situation when a low-level format may be performed on hard disk drives is that when you need to repair a damaged format (parts of the disk become unreadable), or when you wish to wipe away all data on the drive. First of all use the Low Level Formatting utility software provided by the manufacturers of your hard disk drives. Most manufacturers supply low-level format programs for their drives. Here are a few examples:

Seagate: ftp://ftp.seagate.com/techsuppt/seagate_utils/sgatfmt4.zip
IBM: http://www.quantum.com/support/csr/software/csr_software.htm

Western Digital: http://www.wdc.com/service/ftp/wddiag/wd_diag.exe
Maxtor: http://www.maxtor.com/SoftwareDownload/utilities.html
Quantum: http://www.quantum.com/support/csr/software/csr_software.htm

Some of the general purpose Low Level Format Utilities are: Disk Manager by Ontrack and the MicroScope program by Micro 2000. Another excellent, inexpensive, general-purpose PC hardware diagnostic that includes low-level format capability is #1 TuffTEST Pro from #1-PC Diagnostics.

A true LLF program bypasses the system BIOS and sends commands directly to the disk controller hardware. For this reason, many LLF programs are specific to the disk controller hardware for which they are designed. Having a single format program that will run on all types of controllers is virtually impossible. Many hard drives have been incorrectly diagnosed as being defective because the wrong format program was used and the program did not operate properly.

Partitioning

Partitioning a hard disk is the act of defining areas of the disk for an operating system to use as a volume. Each file system has its own method to allocate file space in logical units called *clusters* or *allocation units*. Every hard disk drive must have at least one partition on it. Each file system can support the same or different type of file systems. Three common file systems are used by PC operating systems today:

FAT (File Allocation Table)

FAT is the standard file system supported by DOS, Windows 9x/Me/2000, and Windows NT. FAT partitions support filenames of 11 characters maximum (8 characters + 3-character extension) under DOS, and 255 characters under Windows 9x/Me/2000 or Windows NT 4.0 (or later). The standard FAT file system can support a maximum hard disk drive size of 2GB. Using FDISK, you can create only two physical FAT partitions on a hard disk drive—primary and extended—but you can subdivide the extended partition into as many as 25 logical volumes.
Alternative partitioning programs, such as Partition Magic, can create up to four primary partitions or three primary and one extended.

FAT32 (File Allocation Table, 32-bit)

FAT32 is file system supported by Windows 95 OSR2 (OEM Service Release 2), Windows 98, Windows Me, and Windows 2000. FAT32 can support a maximum single volume size of 2TB or 2,048GB.

NTFS (Windows NT File System)

The native file system for Windows NT/2000 that supports filenames up to 256 characters long and partitions up to (a theoretical) 16 exabytes. NTFS also provides extended attributes and file system security features that do not exist in the FAT file system. Of these three file systems, the FAT file system still is by far the most popular and is accessible by nearly every operating system, which makes it the most compatible as well. FAT32 and NTFS provide additional features but are not universally accessible by other operating systems.

FDISK is used to make the partitions from single megabyte or 1% of the drive up to the full size hard disk drive. Keep in mind the operating system limits). It is recommended to create minimum possible partitions. After a drive is partitioned, each partition must then be high-level formatted by the operating system that will use it.

Warning
FDISK cannot be used to change the size of a partition; all it can do is remove or create partitions. The act of removing or creating a partition destroys any data that was contained in the partition or was on that part of the disk. To manipulate partitions without destroying data, you can use third-party utility programs, such as Partition Magic from PowerQuest or Partition Commander from V Communications.

High-Level Formatting

During the high-level format, the operating system (such as Windows 9x/Me/2000, Windows NT, or DOS) writes the structures necessary for managing files and data on the disk. High-level formatting is not really a physical formatting of the drive, but rather the creation of a table of contents for the disk.

Parts of a Fat Partition (Logical Drive)

FAT partitions have three important parts on each of the logical drive:

Boot Sector (VBS)

This sector contains data that describes the partitions by their starting and ending cylinder, head, and sector locations. The partition table also indicates to the ROM BIOS which of the partitions is bootable and, therefore, where to look for an operating system to load.

File Allocation Table (FAT)

A table held near the outer edge of a disk that tells which sectors are allocated to each file and in what order.

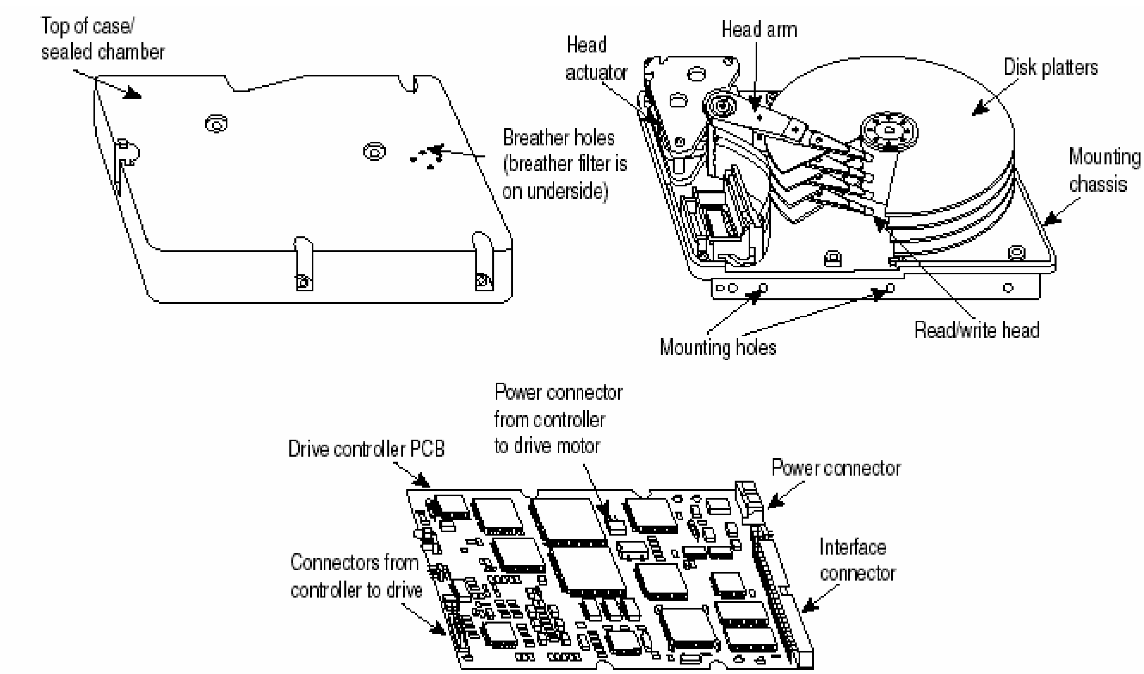
Root Directory

It is a directory in which the operating system stores information about the names and starting locations of the files on the disk.

Main Hard Disk Drive Components

Many types of hard disk drives are on the market, but nearly all share the same basic physical components. Some differences might exist in the implementation of these components (and in the quality of the materials used to make them), but the operational characteristics of most drives are similar. The basic components of a typical hard disk drive are as follows (see Figure 10.6):

- Disk platters
- Read/write heads
- Head actuator mechanism
- Spindle motor (inside platter hub)
- Logic board (controller or Printed Circuit Board)
- Cables and connectors
- Configuration items (such as jumpers or switches)



Typical hard disk drive components

Hard Disk Platters (Disks)

A hard disk drive has one or more platters, or disks. Hard disks for PC systems have been available in a number of form factors over the years. Normally, the physical size of a drive is expressed as the size of the platters. Following are the platter sizes that have been associated with PC hard disk drives:

- 5 1/4-inch
- 3 1/2-inch
- 2 1/2-inch
- 1-inch

Currently, the 3 1/2-inch drives are the most popular for desktop and some portable systems, whereas the 2 1/2-inch and smaller drives are very popular in portable or notebook systems. Most hard disk drives have two or more platters, although some of the smaller drives used in portable systems have only one. The number of platters a drive can have is limited by the drive’s vertical physical size. Platters can be made of auminum/magnesium and glass ceramic composite etc. Modern hard disk drives made by companies such as IBM, Seagate, Toshiba, Areal Technology, and Maxtor currently use glass or glass-ceramic platters.

Recording Media

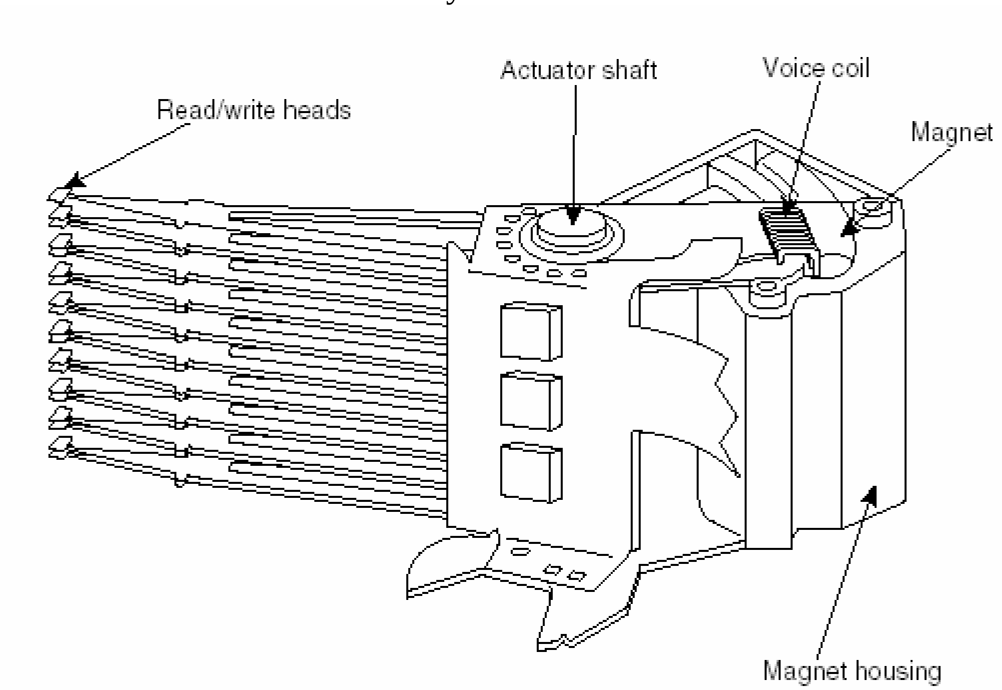
The platters are covered with a thin layer of a magnetically retentive substance, called the *medium*, on which magnetic information is stored. Two popular types of magnetic media are used on hard disk platters:

- Oxide medium
- Thin-film medium

The oxide medium is made of various compounds, containing iron oxide as the active ingredient. The magnetic layer is created on the disk by coating the aluminum platter with a syrup containing ironoxide particles. The thin-film medium is thinner, harder, and more perfectly formed than oxide medium. Floating heights as small as 15nm (nanometers, or about 0.6 µ-inch) above the surface are possible. When the head is closer to the platter, the density of the magnetic flux transitions can be increased to provide greater storage capacity. Additionally, the increased intensity of the magnetic field during a closer-proximity read provides the higher signal amplitudes needed for good signal-to-noise performance. If you open a drive to look at the platters, you would see that platters coated with the thin-film medium look like mirrors.

Read/Write Heads

A hard disk drive usually has one read/write head for each platter surface (meaning that each platter has two sets of read/write heads—one for the top side and one for the bottom side). These heads are connected, or ganged, on a single movement mechanism. The heads, therefore, move across the platters together. The picture below shows a typical hard disk head-actuator assembly.



Read/write heads and rotary actuator assembly.

When the drive is at rest, the heads are forced into direct contact with the platters by spring tension, but when the drive is spinning at full speed, air pressure develops below the heads and lifts them off the surface of the platter. On a drive spinning at full speed, the distance between the heads and the platter can be anywhere from 0.5 to 5 µ-inch or more in a modern drive.

Head Actuator Mechanisms

This mechanism moves the heads across the disk and positions them accurately above the desired cylinder.

Spindle Motors

The motor that spins the platters is called the spindle motor because it is connected to the spindle around which the platters revolve. Spindle motors in hard disk drives are always

connected directly; no belts or gears are involved. The platters in hard disk drives revolve at speeds ranging from 3,600rpm to 15,000rpm (60–250 revolutions per second) or more.

Logic Boards

All hard disk drives have one or more logic boards mounted on them. The logic boards contain the electronics that control the drive's spindle and head actuator systems and present data to the controller in some agreed-upon form. On ATA drives, the boards include the controller itself, whereas SCSI drives include the controller and the SCSI bus adapter circuit. Many disk drive failures occur in the logic board, not in the mechanical assembly. (This statement does not seem logical, but it is true.) Therefore, you sometimes can repair a failed drive by replacing the logic board rather than the entire drive. Replacing the logic board, moreover, enables you to regain access to the data on the drive—something that replacing the entire drive does not provide.

In many cases, logic boards plug into the drive and are easily replaceable. These boards are usually mounted with standard screw hardware. If a drive is failing and you have a spare, you might be able to verify a logic-board failure by taking the board off the known good drive and mounting it on the bad one.

Cables and Connectors

Most of the hard disk drives have at least these three types of connectors:

- Interface connector(s)
- Power connector

Of these, the interface connectors are the most important because they carry the data and command signals between the system and the drive. In most cases, the drive interface cables can be connected in a daisy-chain or bus-type configuration.

The power connector is usually the same four-pin type that is used in floppy disk drives, and the same power-supply connector plugs into it. Most hard disk drives use both 5- and 12-volt power, although some of the smaller drives designed for portable applications use only 5-volt power. In most cases, the 12-volt power runs the spindle motor and head actuator, and the 5-volt power runs the circuitry.

Make sure your power supply can supply adequate power for the hard disk drives installed in your system.

In most computers, the hard disk drive is mounted directly to the chassis using screws, or the drive is grounded via the ground wires in the power connector, so an extra ground wire is unnecessary.

Configuration Items

To configure a hard disk drive for installation in a system, you usually must set several jumpers properly. (See Section 9.9 Dual-Drive Configurations)

Hard Disk Drive Performance

When you want to buy a hard disk, one of the important things to consider is the hard disk performance, which can be measured in terms of:

Average seek time

Average seek time, normally measured in milliseconds (ms), is the average amount of time it takes to move the heads from one cylinder to another a random distance away. One way to measure this specification is to run many random track-seek operations and then divide the timed results by the number of seeks performed. This method provides an average time for a single seek.

Transfer rate

There are different types of transfer rates. You should look for the media transfer rate while looking at the specifications of a hard disk. The media transfer rate is the true rate at which data can be read from the disk, which is how fast data can be read from the drive platters (media). It is the maximum rate that any sustained transfer can hope to achieve.

Hard Disk Capacity

Four figures are commonly used in advertising drive capacity:
Unformatted capacity, in millions of bytes
Formatted capacity, in millions of bytes
Unformatted capacity, in megabytes
Formatted capacity, in megabytes

The term *formatted*, in these figures, refers to the low-level (or physical) formatting of the drive. Most manufacturers of ATA/IDE and SCSI drives now report only the formatted capacities because these drives are delivered preformatted. Usually, advertisements and specifications refer to the unformatted or formatted capacity in millions of bytes because these figures are larger than the same capacity expressed in megabytes.

IDE Interface

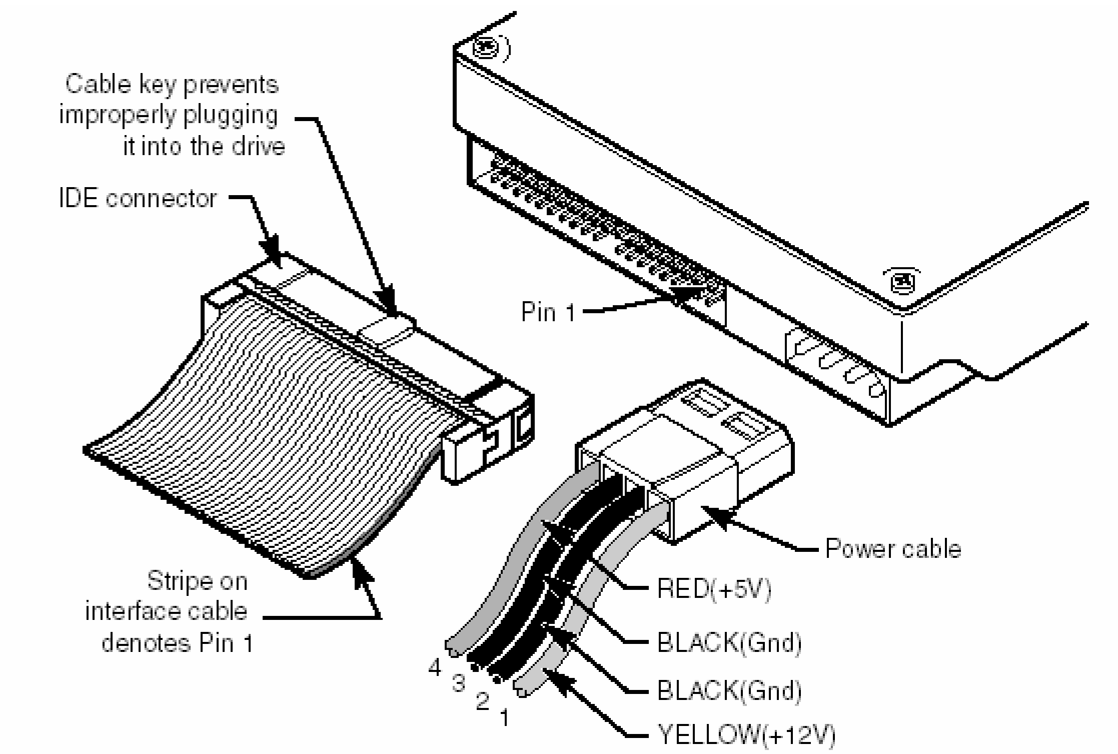
IDE stands for Integrated Drive Electronics. Integrated Drive Electronics refers to the fact that the interface electronics or controller is built into the drive and is not a separate board, as with earlier drive interfaces.
The actual name of this interface is ATA (AT Attachment). This name was given to it as the combined drive and controller were designed to connect to the IBM AT computer. ATA is used to connect not only hard disks, but also CD-ROM and CD-RW drives, DVD drives and tape drives. ATA originally referred to a hard disk drive that plugged directly into a version of the AT-bus, more commonly known as the 16-bit ISA bus. There were many different types of IDE interface variations before the ATA IDE interface was developed with its present specifications. ATA IDE is the interface we will study in the following section.

ATA Features

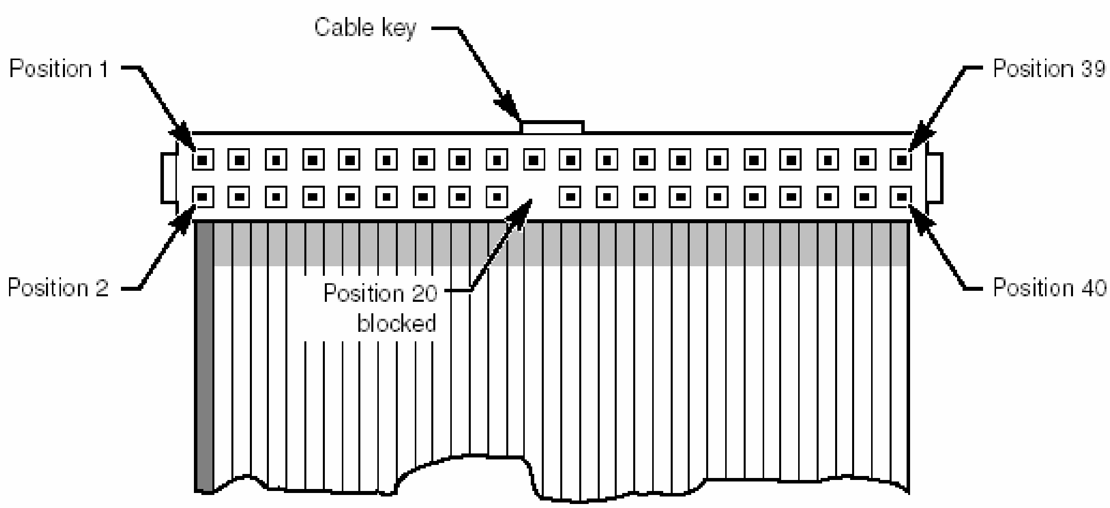
The ATA standards have gone a long way toward eliminating incompatibilities and problems with interfacing IDE drives to ISA/PCI bus systems. The ATA specifications define the signals on the 40-pin connector, the functions and timings of these signals, cable specifications, and so on. The following section lists some of the elements and functions defined by the ATA specification.

ATA I/O Connector

The ATA interface connector is normally a 40-pin header-type connector and is keyed to prevent the possibility of installing it upside down.



Typical ATA (IDE) hard drive connectors.

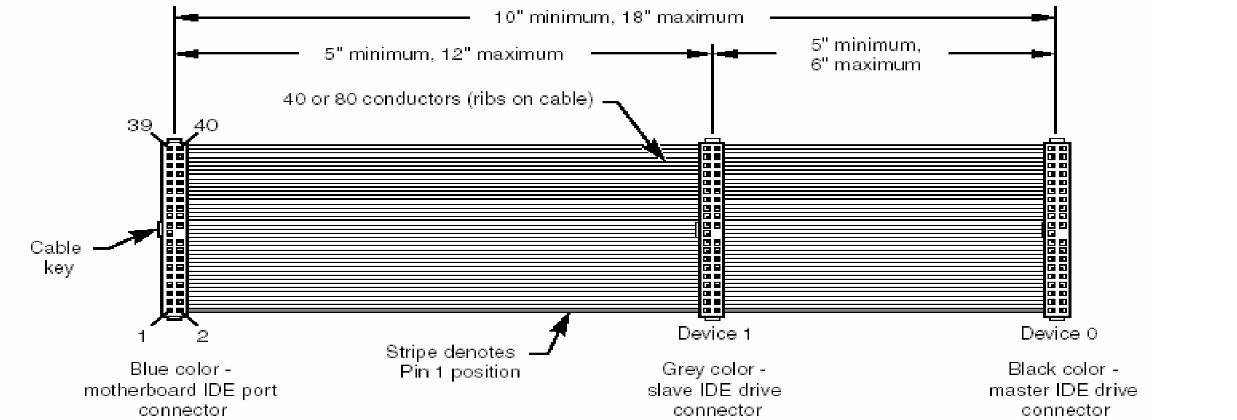


ATA (IDE) 40-pin interface connector detail

To create a keyed connector, the manufacturer generally removes pin 20 from the male connector and blocks pin 20 on the female cable connector, which prevents the user from installing the cable backward. Some cables also incorporate a protrusion on the top of the female cable connector, which fits into a notch in the shroud surrounding the mating male connector on the device. The use of keyed connectors and cables is highly recommended; plugging in an IDE cable backward normally won't cause any permanent damage. However, it can lock up the system and prevent it from running at all.

ATA I/O Cable

A 40-conductor ribbon cable is specified to carry signals between the bus adapter circuits and the drive (controller). To maximize signal integrity and eliminate potential timing and noise problems, the cable should not be longer than 18 inches (0.46 meters). The picture below shows the typical ATA cable layout and dimensions.



ATA (IDE) cable, with 40-pin connectors and either 40- or 80-conductor cables (additional wires are grounded in 80-conductor versions).

Two primary variations of IDE cables are used today: one with 40 conductors and the other with 80 conductors. Both use 40-pin connectors, and the additional wires in the 80-conductor version are simply wired to ground. The additional conductors are designed to reduce noise and interference and are required when setting the interface to run at 66MB/sec (ATA/66) or faster. The drive and host adapter are designed to disable the higher-speed ATA/66 or ATA/100 modes if an 80-conductor cable is not detected. The 80-conductor cable can also be used at lower speeds; although this is unnecessary, it improves the signal integrity.

Dual-Drive Configurations

The ATA standard provides the option of operating on the AT bus with two drives in a daisy-chained configuration. The primary drive (drive 0) is called the master, and the secondary drive (drive 1) is called the slave. You designate a drive as being master or slave by setting a jumper or switch on the drive or by using a special line in the interface called the cable select pin and setting the CS jumper on the drive.

Most IDE drives can be configured with four possible settings:

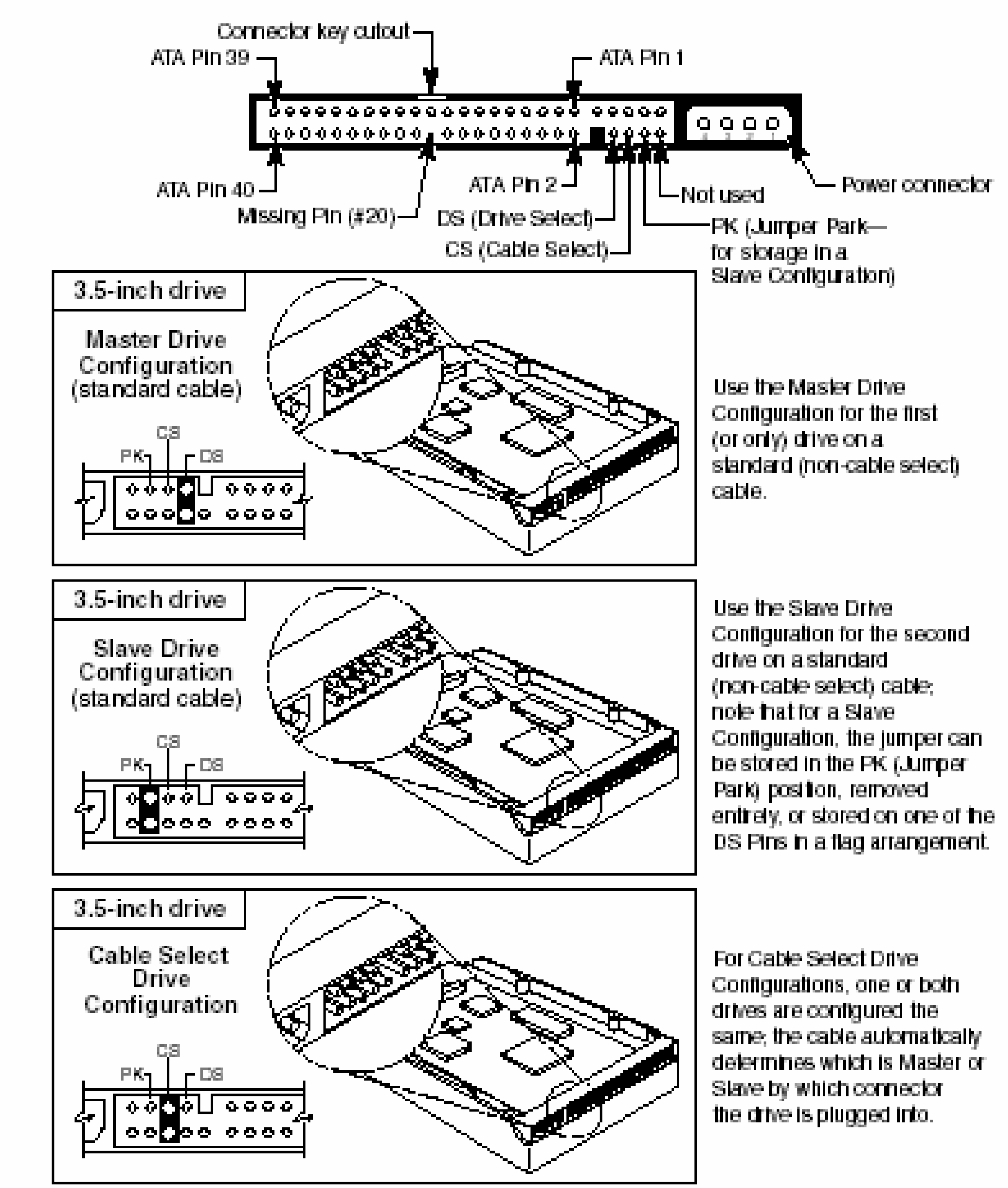
- Master (single-drive)
- Master (dual-drive)
- Slave (dual-drive)
- Cable select

Most of the drives have three settings: master, slave, and cable select. Because each IDE drive has its own controller, you must specifically tell one drive to be the master and the other to be the slave. Telling a drive that it's the slave also usually causes it to delay its spinup for several seconds to allow the master to get going and thus to lessen the load on the system's power supply. Most drives that fully follow the ATA specification now need only one jumper (master/slave) for configuration.

Jumper Settings for Most ATA IDE-Compatible Drives on Standard
(Non–Cable Select) Cables

Jumper Name	Single-Drive	Dual-Drive Master	Dual-Drive Slave
Master (M/S)	On	On	Off
Slave Present (SP)	Off	On	Off
Cable Select (CS)	Off	Off	Off

The master jumper indicates that the drive is a master or a slave. To eliminate confusion over master/slave settings, most newer systems now use the cable select option. Both drives are then configured in cable select mode via the CS jumper on each drive. The picture below shows the jumpers on a typical ATA drive.



ATA (IDE) drive jumpers for most drives. Note that the drive is shown upside down in this figure, whereas previous figures in this chapter depict drives as they appear right side up.

Hard Disk Installation Procedures

To install a hard drive in a PC, you must perform some or all of the following procedures:

- Configure the drive.
- Physically install the drive.

- Configure the system to recognize the drive.
- Partition the drive (FDISK).
- High-level format the drive (FORMAT).

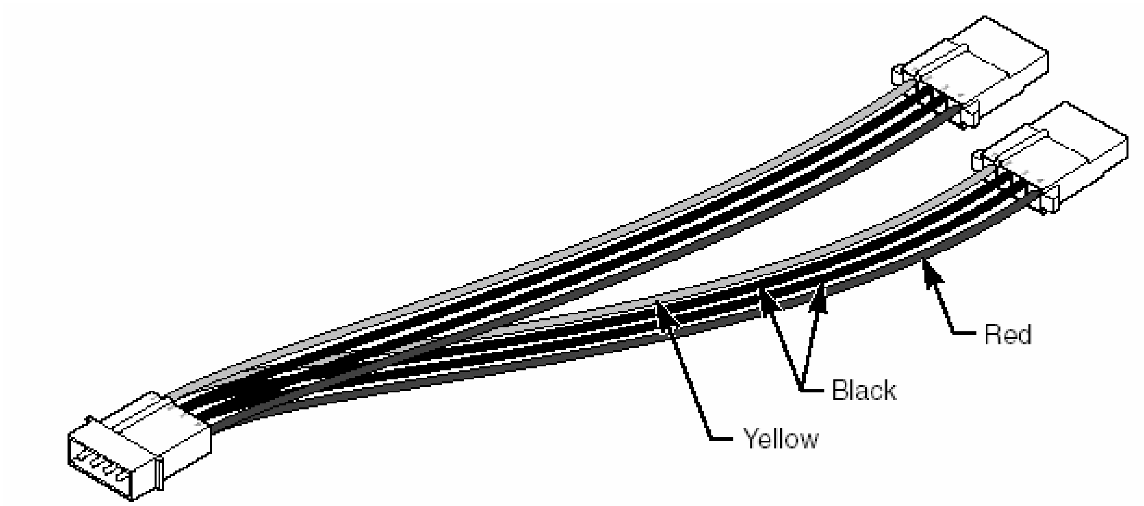
Drive Configuration

Before you physically install a hard disk drive in the computer, you must ensure that it is properly configured. For an ATA/IDE drive, this generally means designating the drive as a master/slave or using the Cable Select (CS) feature and a special cable to determine the relationship.

Physical Installation

The step-by-step procedures for installing a hard drive are as follows:

1. Check your computer for an unused ATA/IDE connector. Typical Pentium-class and newer PCs have two ATA/IDE connectors, allowing for up to four ATA/IDE devices.You might need an additional cable if both master and slave connectors are used on the primary and you are adding a third device.
2. Double-check the pin configuration and cable type. The colored (normally red or red-flecked) stripe on one edge of the cable goes to pin 1 of the hard drive's data connector. Most cables and drive connectors are keyed to prevent improper (backward) installation, but many are not. Note that pin 1 on the drive connector is almost always oriented nearest to the power connector.
3. Configure the drive jumpers. If the drive is ATA/IDE and you are using a cable that supports CS, you must set the CS jumper on any drives connected to that cable. Otherwise, you must set the drives on the cable as either master or slave.
4. Slide the drive carefully into a drive bay of the correct size. Most hard drives are 3 1/2- inch wide and one-inch high. Be sure the screws aren't too long; if you bottom the screws in the drive, you can damage it.
5. Attach the data cable connector to the back of the drive, unless you are adding the drive and cable at the same time.
6. Attach the appropriate power connector to the drive. Most hard drives use the larger, or "Molex," four-wire power connector. If necessary, you can use a Y-splitter cable to create two power connectors from one (many computers have fewer power connectors than drive bays).



Power cord splitter and connector

7. Turn on the computer and listen for the new hard disk to spin up. Even though today's drives are very quiet compared to early hard disks, you should still hear a faint rattling or clicking as the new hard disk starts to run. If you don't hear anything from the drive, double-check the data and power cables.
8. Restart the computer and access the BIOS setup screens to configure the new hard disk. At a minimum, you'll need to detect (or set) the following parameters: cylinders, heads, sectors per track (called CHS parameters), and write precompensation. If your BIOS has an autodetect or auto type setting, because it will configure most parameters automatically. For IDE hard drives above 528MB (504 megabytes), you also must set logical block address (LBA) translation to access the drive's full capacity. If autodetect was selected, LBA translation should already be set correctly. Many systems have a Peripherals Configuration screen, which also enables you to set transfer speeds for maximum drive performance, but again these should be configured automatically if you enabled autodetect. If your BIOS cannot autodetect your drive, see the documentation that came with the drive for the correct settings. Most drives have the CHS parameters as well as jumper setting information displayed on a sticker attached to the drive, and all modern BIOSes support autodetect, as well. Save the BIOS configuration and exit the BIOS Setup screen to continue. The following section has more information on this step, as it can be complicated.
9. Restart the computer and prepare to run FDISK to prepare the hard drive for formatting and use. Or, you can use drive partitioning software, such as PowerQuest's PartitionMagic, to automatically create and manage disk partitions. Normally, FDISK can be found on the operating system start disk, such as the Windows startup disk. Simply insert the startup disk in the floppy drive, boot from it, and type the FDISK command at the A: prompt.

System Configuration

After your drive is physically installed, you must provide the computer with basic information about the drive so the system can access and boot from it. How you set and store this information depends on the type of drive and system you have. Standard (IDE) setup procedures apply for most hard disks.

Formatting

Proper setup and formatting are critical to a drive's performance and reliability. This section describes the procedures used to format a hard disk drive correctly. Use these procedures when you install a new drive in a system or immediately after you recover data from a hard disk that has been exhibiting problems. The three major steps in the formatting process for a hard disk drive subsystem are as follows:

- Low-level formatting
- Partitioning
- High-level formatting

Drive Partitioning with FDISK

Partitioning a hard disk is the act of defining areas of the disk for an operating system to use as a volume. The FDISK program is the accepted standard for partitioning hard disk drives for use with all operating systems. The program is included with your operating system, and although it has the same name and basic functions with any OS, you should normally use the one that specifically came with your OS. Partitioning prepares the boot sector of the disk in such a way that the FORMAT.COM program or the Windows GUI format utility can operate correctly. FDISK also makes it possible for various operating systems to coexist on a single hard disk. No matter which operating system you use, it should come with an FDISK program that can be used to partition the drive.

Note
Because FDISK depends on BIOS information about the hard disk to determine the size and drive geometry of the hard disk, having correct BIOS settings saved in the BIOS Setup is vital to the correct operation of FDISK. If a 10GB hard drive is defined in the BIOS as a 10MB hard drive, for example, all that FDISK will prepare for use—and all that FORMAT will ultimately prepare—is 10MB.

Primary and Extended Partitions

With any version of Windows, as with MS-DOS, FDISK enables you to create two types of disk partitions: primary and extended. A primary partition can be bootable but an extended partition cannot. If you have only a single hard disk in your system, at least part of the drive must be prepared as a primary partition if you want to start your computer from the hard disk. A primary partition is seen as a single volume or drive letter (C: on one-drive systems), whereas an extended partition acts as a sort of logical container for additional volumes (drive letters D: and beyond). A single extended partition can contain a single volume (also referred to by FDISK as a *logical DOS drive*) or several volumes (logical DOS drives) of various sizes.

Don't get hung up on the fact that FDISK calls partitions "DOS" partitions or "DOS" drives. This is true even though the operating system you are installing is Windows 95, 98, Me, NT, 2000 and so on. Depending on the version of Windows in use (and with any version of MS-DOS), you might need to subdivide a hard drive through the use of FDISK. The original release of Windows 95 and MS-DOS supports a file system known as FAT16, which allows no more than 65,536 files per drive and a single drive letter or volume size of no more than 2.1GB in size. Thus, a 10.1GB hard disk prepared with MS-DOS or the original Windows 95 (or 95A) must have a minimum of five drive letters, and could have more.
Another reason for subdividing a single drive into multiple volumes is for increased data security. For example, PowerQuest (the creator of PartitionMagic) suggests a three-volume partitioning scheme that looks like this:

C: for the operating system and utilities
D: for applications
E: and above for data

In this example, primary and extended partitions would be assigned as shown here:

C: the primary partition
D: and E: in the extended partition as logical DOS drives (volumes)

If a catastrophic failure to the disk structure wipes out C: or D:, drive E:—which contains the data—will normally still be intact. This method also makes backing up the data easier; just set the backup program to back up all of E:, or all the changed files on E:.

FDISKcan be used in many ways, depending on the number of hard drives you have in your system and the number of drive letters you want to create. With a single drive, creating a primary partition (C:) and an extended partition with two logical DOS drives within it results in the following drive letters:

Partition Type	**Drive letter(s)**
Primary C:	Extended D: and E:

Running FDISK

When you run newer versions of FDISK, the first thing that happens is you are prompted with the following:

Do you wish to enable large disk support (Y/N)...........? [Y]

If you answer 'Yes' to this question, FDISK creates FAT32 volumes for all volumes created that are larger than 512MB. Answering No to this question forces FDISK to create only FAT16 volumes, which are limited to a 2GB maximum size and waste more space on the disk due to larger cluster sizes. Normally, in a modern system you would answer Yes, allowing the use of FAT32. After answering the question, FDISK shows a menu similar to the following:

Current fixed disk drive: 1

Choose one of the following:

1. Create DOS partition or Logical DOS Drive
2. Set active partition
3. Delete partition or Logical DOS Drive
4. Display partition information
5. Change current fixed disk drive

Enter choice: [1]

Option 5 is shown only if FDISK detects more than one drive on your system (if more than one is entered via your BIOS Setup). In that case, FDISK defaults to the first drive, and via option 5 you can cause FDISK to work with any of the other hard disks on the system. To create partitions, you select option 1. If the drive is already partitioned, however, you can use option 4 to display the current layout of the drive. After selecting

option 1, the menu changes to enable you to create primary or extended partitions on a drive as follows:

Create DOS Partition or Logical DOS Drive

Current fixed disk drive: 1

Choose one of the following:
1. Create Primary DOS Partition
2. Create Extended DOS Partition
3. Create Logical DOS Drive(s) in the Extended DOS Partition

Enter choice: [1]

The rules require you to create a primary partition first on the boot drive, but on a secondary or non-boot drive, you can create just an extended partition if you choose. So, if you are partitioning the first drive in a system and it will be bootable, you would choose option 1.

At this point, you are prompted to decide whether you want to use the maximum available size for a Primary DOS Partition. If you say ‘Yes’ and you’re using FAT32, the primary partition uses the entire drive. Conversely, if you say “No” and you did not enable large drive support (meaning you are using FAT16), the partition uses the entire drive or 2GB, whichever is smaller.

If you decide to create a primary partition that is not the full drive, you should go back through the menus and create an extended partition using the rest of the drive, and then further divide it into logical drives. After all the partitions are created, the final operation is to make one of them active (bootable), which is option 2 from the main FDISK menu. Normally, the only one that can be active is the primary partition.

After that is done, all FDISK operations are complete, and you can exit the program. When exiting FDISK after making partition changes, the system must be rebooted before these changes will be recognized. After rebooting, you must then high-level format each of the volumes with the operating system FORMAT command, which then allows the operating system to store files on them.
The C: volume must normally be formatted with the system files, although when you install Windows via the Windows Setup command, it detects whether the system files are present and offers to install them for you.

High-Level (Operating System) Formatting

The final step in the installation of a hard disk drive is the high-level format. Similar to the partitioning process, the high-level format is specific to the file system you've chosen to use on the drive. On Windows 9x and DOS systems, the primary function of the high-level format is to create a FAT and a directory system on the disk so the operating system can manage files. You must run FDISK before formatting a drive. Each drive letter created by FDISK must be formatted before it can be used for data storage.

Usually, you perform the high-level format with the FORMAT.COM program or the formatting utility in Windows 9x Explorer. FORMAT.COM uses the following syntax: FORMAT C: /S /V

This command high-level formats drive C:, writes the hidden operating system files in the first part of the partition, and prompts for the entry of a volume label to be stored on the disk at the completion of the process.

1. The FAT high-level format program performs the following functions and procedures: Scans the disk (read-only) for tracks and sectors marked as bad during the LLF, and notes these tracks as being unreadable.
2. Returns the drive heads to the first cylinder of the partition, and at that cylinder (Head 1, Sector 1) writes a DOS volume boot sector.
3. Writes a FAT at Head 1, Sector 2. Immediately after this FAT, it writes a second copy of the FAT. These FATs are essentially blank except for bad-cluster marks noting areas of the disk that were found to be unreadable during the marked-defect scan.
4. Writes a blank root directory.
5. If the /S parameter is specified, copies the system files, IO.SYS and MSDOS.SYS (or IBMBIO.COM and IBMDOS.COM, depending on which DOS you run), and COMMAND.COM to the disk (in that order).
6. If the /V parameter is specified, prompts the user for a volume label, which is written as the fourth file entry in the root directory.

Now, the operating system can use the disk for storing and retrieving files, and the disk is a bootable disk. During the first phase of the high-level format, the program performs a marked defect scan. Defects marked by the LLF operation show up during this scan as being unreadable tracks or sectors. When the high-level format encounters one of these areas, it automatically performs up to five retries to read these tracks or sectors. If the unreadable area was marked by the LLF, the read fails on all attempts.
After five retries, the DOS FORMAT program gives up on this track or sector and moves to the next one. If an area remains unreadable after the initial read and five retries, it is marked in the FAT as a bad cluster.

Note
Because the high-level format doesn't overwrite data areas beyond the root directory of the hard disk, using programs such as Norton Utilities to unformat the hard disk that contains data from previous operations is possible, provided no programs or data has been copied to the drive after high-level formatting. Unformatting can be performed because the data from the drive's previous use is still present.

If you create an extended partition, the logical DOS drive letters found there need a simpler FORMAT command because system files aren't necessary—for example, FORMAT D:/V for drive D:, FORMAT E:/V for drive E:, and so on.

Replacing an Existing Drive

Although formatting and partitioning a new hard disk can be challenging, replacing an existing drive\and moving your programs and files to it can be a lot more challenging.

Drive Migration for MS-DOS Users

When MS-DOS 6.x was dominant, many users used this straightforward method to transfer the contents of their old hard drive to their new hard drive:

1. The user creates a bootable disk containing FDISK, FORMAT, and XCOPY.
2. The new hard drive is prepared with a primary partition (and possibly an extended partition, depending on the user's desires).
3. The new hard drive is formatted with system files, although the operating system identifies it as D:.
4. The XCOPY command is used to transfer all non-hidden files from C:\ (the old hard drive) to
   D:\, thus:
   XCOPY C:\ D:\/S/E

The XCOPY command also is used as necessary to transfer files from any remaining drive letters on the old hard drive to the corresponding drive letters on the new drive.
Because the only hidden files such a system would have were probably the operating system boot files (already installed) and the Windows 3.1 permanent swap file (which could be re-created after restarting Windows), this "free" data transfer routine worked well for many people.

After the original drive was removed from the system, the new drive would be jumpered as master and assigned C:. The user then would need to run FDISK from a floppy and set

the primary partition on the new C: drive as Active. Then, the user would exit FDISK and the drive would boot.

Drive Migration for Windows 9x/Me Users

Windows 9x/Me have complicated the once-simple act of data transfer to a new system by their frequent use of hidden files and folders (such as \Windows\Inf, where Windows 9x hardware drivers are stored). The extensive use of hidden files was a major reason for a greatly enhanced version of XCOPY to be included in Windows 9x/Me.

> **Note**
> XCOPY32 is automatically used in place of XCOPY when XCOPY is started within a DOS session under Windows.

XCOPY32 for Windows 9x Data Transfer

Compared to "classic" XCOPY, XCOPY32 can copy hidden files; can preserve file attributes such as system, hidden, read-only, and archive; can automatically create folders; and is compatible with long filenames. Thus, using it to duplicate an existing drive is possible, but with these cautions:

1. The XCOPY32 command is much more complex.
2. Errors might occur during the copy process because of Windows's use of temporary files during normal operation, but XCOPY32 can be forced to continue.

This command line calls XCOPY32 and transfers all files and folders with their original attributes intact from the original drive (C:) to the new drive (D:). This command, however, must be run from an MS-DOS prompt window (and not MS-DOS Mode) under Windows 9x/Me, as follows:

```
xcopy32 c:\. d:\ /e/c/h/r/k
```

The command switches are explained here:

- /e. Copies folders, even if empty; also copies all folders beneath the starting folder.
- /c. Continues to copy after errors. The Windows swap file can't be copied due to being in use.
- /h. Copies hidden and system files.
- /r. Overwrites read-only files.
- /k. Preserves file attributes.

Repeat the command with appropriate drive-letter changes for any additional drive letters on your old drive. After the original drive is removed from the system, the new drive

must be jumpered as master (or single), and the operating system assigns it C:. You next need to run FDISK from a floppy and set the primary partition on the new C: drive as Active. Then, exit FDISK, and the drive will boot.
Note that although the XCOPY method has worked for me, some people have problems with it. A much more automated and easy approach to cloning drives is to use commercial software designed for that purpose, such as Drive Copy by PowerQuest or Norton Ghost by Symantec. Even though many users have prepared new hard drives for use with nothing but FDISK and FORMAT, today's more complex systems are presenting increasingly good reasons for looking at alternatives.

Hard Disk Drive Troubleshooting and Repair

If there is a mechanical defect in the hard disk you could do it but it might not be feasible. If the failure is in the logic board, that board can be replaced with one from another drive which has some other problem. Most hard disk drive problems are not mechanical hardware problems; instead, they are "soft" problems that can be solved by a new LLF and defect-mapping session. Soft problems are characterized by a drive that sounds normal but produces various read and write errors. Hard problems are mechanical, such as when the drive sounds as though it contains loose marbles.

Constant scraping and grinding noises from the drive, with no reading or writing capability, also qualify as hard errors. In these cases, an LLF is unlikely to put the drive back into service. If a hardware problem is indicated, first replace the logic-board assembly.

Testing a Drive

A simple test can be performed to check if your drive has been partitioned and formatted properly. Attach the drive to your system. After detecting the drive in the BIOS and saving the changes, start your operating system from the boot disk.
Then, from the A: prompt, enter the following command: DIR C:
This produces one of the following responses:

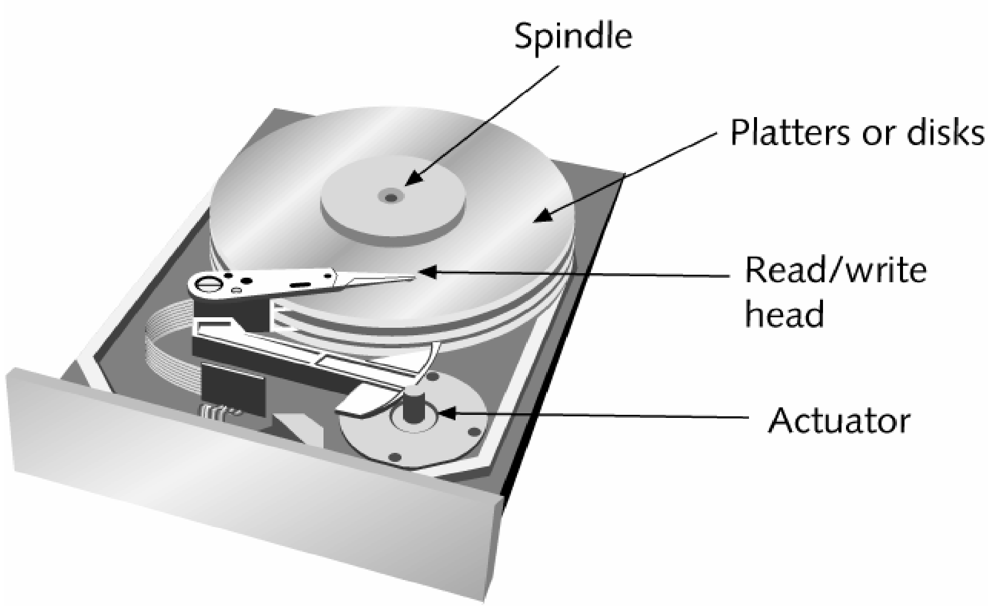
1. Invalid drive specification. This indicates the drive does not have a valid partition (created by FDISK) or that the existing Master Boot Sector or partition tables have been damaged. No matter what, the drive must be partitioned and FORMATted before use. You also get this warning on a FAT32 or NTFS partitioned drive if you use a Windows 95 (original version) or MS-DOS boot disk when checking. Use a Windows 95B, Windows 98/Me, or Windows 2000 boot disk to avoid this false message from FAT32 partitions. Or, use a Windows NT or Windows 2000 boot disk to detect NTFS partitions.
2. Invalid media type. This drive has been partitioned but not FORMATted, or the format has been corrupted. You should use FDISK's #4 option to examine the

drive's existing partitions, and either delete them and create new ones or keep the existing partitions and run FORMAT on each drive letter.

3. Directory of C:. The contents of C: drive are listed, indicating the drive was stored with a valid FDISK and FORMAT structure and data.

Advanced Hard Disk Drives

In this lesson, we broaden our discussion of hard disk drives to include the newer large-capacity drives and cover several of the newest methods.



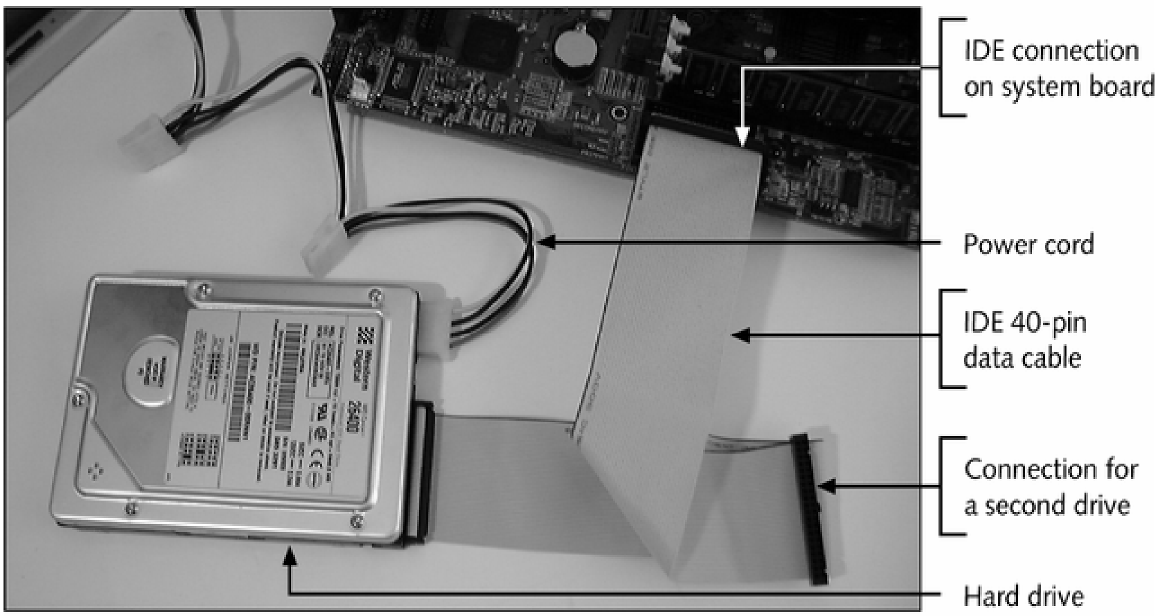
Limitations of Early Hard Disk Drives

The original basic hard disk drives specifically the ST-506—were relatively simple to install because the primary input/output commands were handled by the PC AT system BIOS. They used the routines built into the original IBM AT and the same interface command set as the original ST-506 hard drive. As drive capacities grew, they required many changes in setup to get around the limitations imposed by earlier models. Often, these changes added to the workload of the processor, which had a net effect of slowing down the processing of data. The result was a search for new methods to overcome those bottlenecks, which in turn led to other design considerations that had to be addressed. Storage technology is still evolving. Drives increase in capacity and speed, causing changes in PC design and operating-system support to take advantage of the larger-capacity, faster drives like 80 GB 7200 RPM Western Digital or 40 GB 7200 RPM Seagate or 5200 RPM larger disks.

IDE and EIDE Drives

IDE (Integrated Drive Electronics) drives have been in use since the late 1980s. The purpose of the IDE was to integrate the drive controller with the drive itself rather than

use a separate controller card. The ATA (Advanced Technology Attachment—the official name for IDE drives) standard is based on the original IBM AT standard for hard disk drives. ATA drives use the same interface command set as the original ST-506 drives and are handled by the system BIOS built in to the original IBM AT. ATA was, and is, a good command set, but its limitations led to its decline as a viable hard drive interface. These limitations set the stage for the development of the Enhanced Integrated Drive Electronics (EIDE). The EIDE drive system was developed with two essential objectives: increasing the size of available disk drives and increasing the speed of data transfer between the host and the disk drive.



The EIDE specification:

- Increased the numbers of drives available to the average computer.
- Increased the data transfer rate.
- Allowed for non-hard disk drives such as CD-ROM, ZIP, and tape drives to be configured to EIDE standards and used from an EIDE controller.
- Broke the 528-MB storage capacity limit of the ATA standard.

Chapter 9

High Capacity Disk Drives

OBJECTIVES

CD-ROM Drives

SCSI Drives

CD-ROM Drives

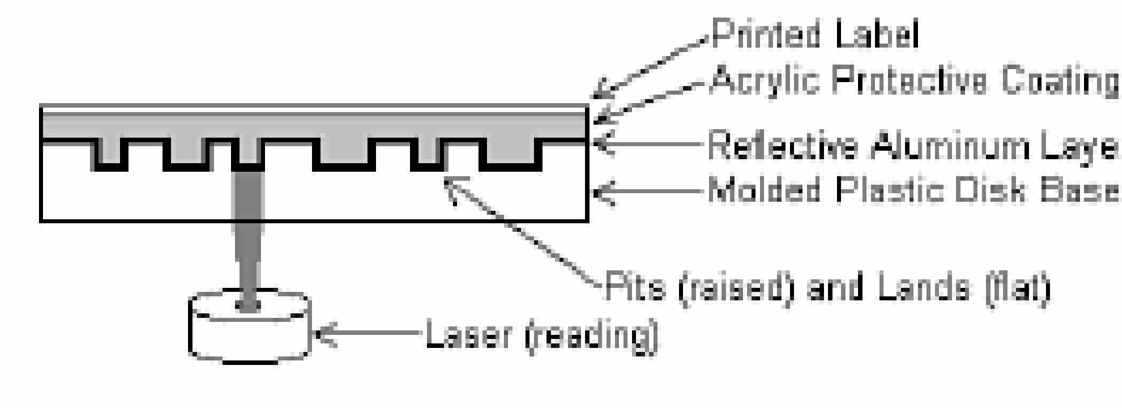
The CD-ROM (compact disc read-only memory) is a technology taken directly from the audio world that has become standard equipment for computers. This lesson covers the basics of installing and using CD-ROM drives. CD-ROM stands for *compact disc read-only memory. This is an optical and read only storage device.* The CD-ROM is a read-only optical storage medium. It is capable of holding up to 650 MB to 700 MB of audio, data or video. Its normal size is a 120mm diameter, 1.2mm thick. It is made up of a plastic disc. CD-R (CD-recordable) and CD-RW (CD-rewritable) and DVD (digital versatile disc) are other variations of the compact disc.

CD-ROM Construction

The term *CD-ROM* refers to both the discs themselves and the drive that reads them. A CD is made of a polycarbonate wafer, 120mm in diameter and 1.2mm thick, with a 15mm hole in the center. This wafer base is stamped or molded with a single physical track in a spiral configuration starting from the inside of the disc and spiraling outward. The track has a pitch, or spiral separation, of 1.6 microns (millionths of a meter, or thousandths of a millimeter).

The laser used to read the disc would pass right through the clear plastic, so the stamped surface is coated with a reflective layer of metal (usually aluminum) to make it reflective. Then, the aluminum is coated with a thin protective layer of acrylic lacquer, and finally a label or printing is added.

Note
CD-ROM media should be handled with the same care as a photographic negative. The CD-ROM is an optical device and degrades as its optical surface becomes dirty or scratched. Also it is important to note that, although discs are read from the bottom, the layer containing the track is actually much closer to the top of the disc. Writing on the top surface of a disc with a ballpoint pen, for example, easily damages the recording underneath. You need to be careful even when using a marker to write on the disc. The inks and solvents used in some markers can damage the print and lacquer overcoat on the top of the disc, and subsequently the information layer right below. Use only markers designed for writing on CDs. The important thing is to treat both sides of the disc carefully, especially the top (label) side.



The construction of a CD-ROM

CD-ROM Speed

CD-ROM drives have been available in speeds from 1x up to 56x and beyond. Most non-rewritable drives up to 12x were CLV; most drives from 16x and up are CAV. With CAV drives, the track data is moving past the read laser at various speeds, depending on where the data is physically located on the CD (near the inner or outer part of the track). This also means that CAV drives read the data at the edge of the disk more quickly than data near the center. This allowed for some misleading advertising when these drives first came out. For example, a 12x CLV drive reads data at 1.84MB/sec no matter where that data is on the disc. On the other hand, a 16x CAV drive reads data at speeds up to 16x (2.46MB/sec) on the outer part of the disc, but it also reads at a much lower speed of only 6.9x (1.06MB/sec) when reading the inner part of the disc. On average, this would be only 11.5x, or about 1.76MB/sec. The average is actually overly optimistic because discs are read from the inside (slower part) out, and an average would relate only to reading completely full discs. The real-world average could be much less than that.

What this all means is that the 12x CLV drive would be noticeably faster than the 16x drive, and faster than even a 20x drive! Remember that all advertised speeds on CAV drives are only the maximum transfer speed the drive can achieve, and it can achieve that only when reading the very end part of the disc.

Adding CDs to DOS/Win3.x

The Microsoft CD Extensions for DOS enable the DOS operating system (and by extension, Windows 3.x) to identify and use data from CD-ROMs attached to the system. The original DOS operating system had no provisions for this technology, so “hooks” or handling of this unique media are not part of the basic operating environment. Using these extensions is convenient for all involved, however.
As CD-ROM technology changes, the MSCDEX can be changed, independently of DOS. For example, most Photo-CD, multiple-session CD-ROM drives require MSCDEX.EXE

version 2.21 or higher, which has been modified from earlier versions to accommodate the newer CD-ROM format. MSCDEX.EXE should be in your software kit with your drive. If not, you can obtain the latest copy directly from Microsoft. If you are a registered user of DOS or Windows 3.1, MSCDEX.EXE is free. Read the licensing agreement that appears on the disc or in your manual concerning the proper licensing of the MSCDEX files.

Your installation software should add a line similar to the following to your AUTOEXEC.BAT file: C:\WINDOWS\COMMAND\MSCDEX.EXE /d:mscd001 C:\WINDOWS\COMMAND is the directory in which the MSCDEX.EXE file is located by default with Windows 9x and later; with MS-DOS and Windows 3.1, it could be located in the \DOS directory or in the directory containing the CD-ROM drivers. The /d:mscd001 portion of the command line supplies the MSCDEX extension with the device signature defined in the CD-ROM device driver of your
CONFIG.SYS file.

Creating a Bootable Floppy with CD-ROM Support

To create a bootable CD, you need a bootable floppy that contains the drivers to support your CD drive in DOS mode (sometimes called real-mode drivers).

Note
Optionally, you can use a Windows 98/Me startup floppy because these have the DOS-level CD-ROM support already configured and installed. You can even use a Windows 98/Me startup disk to boot a Windows 95 system, so getting a disk from a 98/Me system is definitely the easiest way to proceed. If that is not an option, you can add the CD-ROM support to the Windows 95 or any DOS startup disk.

Test your boot floppy (with CD-ROM drivers) by first booting to the floppy. Then, with a CD containing files in the CD-ROM drive, see whether you can change to the CD-ROM drive and read a directory of the files (try the DIR command). The CD usually is the next drive letter after your last hard drive letter. For example, if your last hard drive letter is C:, the CD-ROM will be D:.
If you can display a directory listing of the CD after booting from the floppy, your CD-ROM drivers are properly loaded.

Creating a Rescue CD

A number of programs on the market today allow you to make a compressed image file of the contents of any drive. These programs, such as the Ghost program sold by Symantec or PowerQuest’s Drive Image, enable you to lock in the condition of any drive

as of a particular time. This enables you to create an image file of your system when it's working and use the image-restore feature to reset your system when it fails.

The perfect place to store a compressed image file is on a CD-R. At a minimum, your rescue disc should contain the compressed image file (a 737MB, 80-minute CD-R/RW could contain the equivalent of a nearly 1.5GB drive's normal contents if maximum compression option is used). It's also desirable to place a copy of the image-restore program on the CD. Mastering this type of rescue CD is done exactly the same as a conventional CD mastering process. To use the rescue CD, you must boot your system with drivers that allow the CD-ROM drive to work, run the restore program to read the data from the CD, and overwrite the drive's existing contents.

If you're looking for a single-CD solution to rescuing your system, one that won't require you to lug around a bootable floppy disk, you can burn a rescue CD that is bootable all by itself.

Making a Bootable CD for Emergencies

A little-known capability to PC users is that they can create their own versions of what is standard with more and more new computers: a bootable CD that can be used to start up a system and restore it to a previously saved state.

Files Needed for a Bootable CD

The minimum requirements for a bootable CD include

- A system in which the CD/DVD can be designated as a boot drive
- A CD-R or CD-RW drive and either CD-R or CD-RW media
- Recording software that allows creation of a bootable CD
- A floppy disk containing your operating system boot files

The basic procedure for creating a bootable CD is as follows:

1. Create a bootable floppy for the operating system you want to install on the CD.
2. Get a blank CD-R, and place it into your CD-R/CD-RW drive.
3. Start your mastering software.
4. Under disc layout, ensure that ISO 9660 is selected. This CD format doesn't permit long filenames, so be sure that any additional files you add to the CD have no more than eight characters plus up to three characters for the filename extension.
5. Make sure the bootable option is enabled in the disc layout.
6. When prompted, insert the floppy disk containing boot files into A: drive.

7. These files are copied to your CD layout. Note that the names of these files are not the same as the normal operating system boot files. The files are called BOOTCAT.BIN and BOOTIMG.BIN.
8. Add any additional files (image files, operating system install files, and so on) you want to the layout.
9. Start the CD creation process.
10. When the process is completed, view the contents of the finished CD-R.
11. Close your mastering program, saving the layout if you desire.
12. Insert the bootable CD you just created into your CD-ROM drive.
13. Restart the computer and see whether your system boots from the CD.

If you follow these directions, you will discover that the CD-ROM disc appears as both drive A: and another letter. A: is the floppy image you copied during the creation process, whereas all the other files show up on the main CD drive letter.

CD-Writeable

There are two main types of recordable CD drives and discs are available. One of them is called CD-R (recordable) and the other is CD-RW (rewritable). All the CD-RW drives can also function as CD-R drives. All drives sold today are CD-RW. Those drives can work with either CD-R or CD-RW discs.

Once a CD-R is written with data, it is permanently stored and can’t be erased. The write-once limitation makes this type of disc less than ideal for system backups or other purposes in which it would be preferable to reuse the same media over and over. However, because of the low cost of CD-R media, you might find that making permanent backups to essentially disposable CD-R discs is as economically feasible as tape or other media. In a CD-R the writing process is completed by burning an organic dye.

CD-RW discs can be reused up to 1,000 times, making them suitable for almost any type of data storage task. Most recordable CD drives are both CD-R and CD-RW in one. In a CD-RW the writing process is completed by phase change of a metal alloy. The writing speed of a CD-RW is less than a CD-R.

Warning
Do not try to write a CD-R or CD-RW more than the manufacturer’s rated speed.

The capacity of a CD-R or CD-RW is normally 650 MB to 700 MB.

CD-R Color

The following is a comparison of different colors of the compact disc:

Gold-gold	Less tolerance for power variations Works best in drives that use a Long Write Strategy (longer laser pulse) to mark media
Gold-green	Has a rated life span of 10 years Works best in drives that use a Short Write Strategy (shorter laser pulse) to mark media
Silver-blue	Similar performance to green media plus rated to last up to 100 years A good choice for long-term archiving

Selecting CD-Writers

There are five major factors influence your ability to create a working CD-R:

- Interface type
- Drive buffer size
- The location and condition of the data you want to record to CD-R.
- The recording speed
- Whether the computer is performing other tasks while trying to create the CD-R.

To improve the chances of getting reliable CD-R/RW creation, look for drives with

- *A large data buffer (2MB or larger), or better yet some form of buffer under-run protection.* The data buffer in the drive holds information read from the original data source, so that if a pause in data reading occurs, there's less of a possibility of a buffer under-run until the on-drive buffer runs empty. A larger buffer minimizes the chance of "running out of data." Newer drives with buffer under-run protection virtually eliminate this problem, no matter what size buffer is in the drive.

- *Support for UDMA operating modes.* UDMA modes transfer data more\quickly and with less CPU intervention than earlier versions of ATA.

Warning
Also, if you have problems with reliable CD-R creation at the drive's maximum speed, try using a lower speed (4x instead of 8x, for example). Your mastering job will take twice as long, but it's better to create a working CD-R slowly than ruin a blank quickly.

Packet Writing

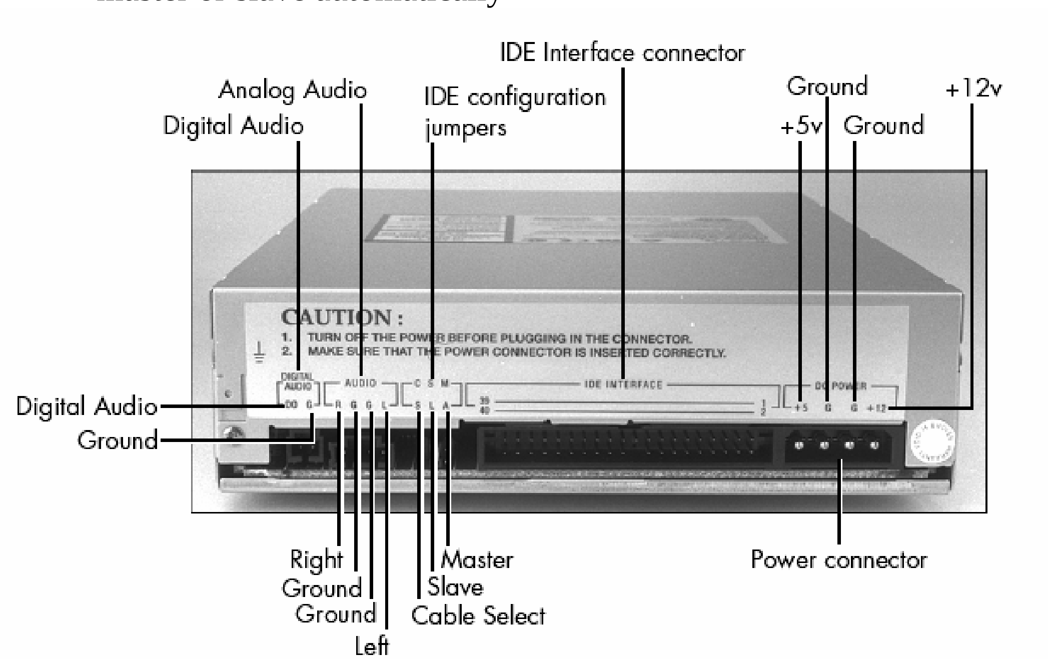
An alternative approach to above is to use packet-writing software to create your CD-R. All late-model CDR/CD-RW drives support packet writing, which allows drag-and-drop copying of individual files to the CD-R/RW rather than transferring all the files at once as with normal mastering software. This “a little at a time” approach means that less data must be handled in each write and can make the difference between success and failure. If your drive supports this feature, it probably includes packet-writing software in the package. Note that although packet-written CDs can be read with Windows 9x, Me, NT, and 2000, they can’t be read with Windows 3.1 and MS-DOS because these operating systems don’t have drivers available that support packet-written CDs.

Installing an Optical Drive

Drive Configuration

Configuration of your new optical drive is paramount to its proper functioning. Examine your new drive (see Figure 11.2) and locate any jumpers. For an IDE drive, here are the typical ways to jumper the drive:

1. As the primary (master) drive on the secondary IDE connection
2. As the secondary (slave) drive to a current hard disk drive
3. Using Cable Select (CS), where the cable connector determines which drive is master or slave automatically

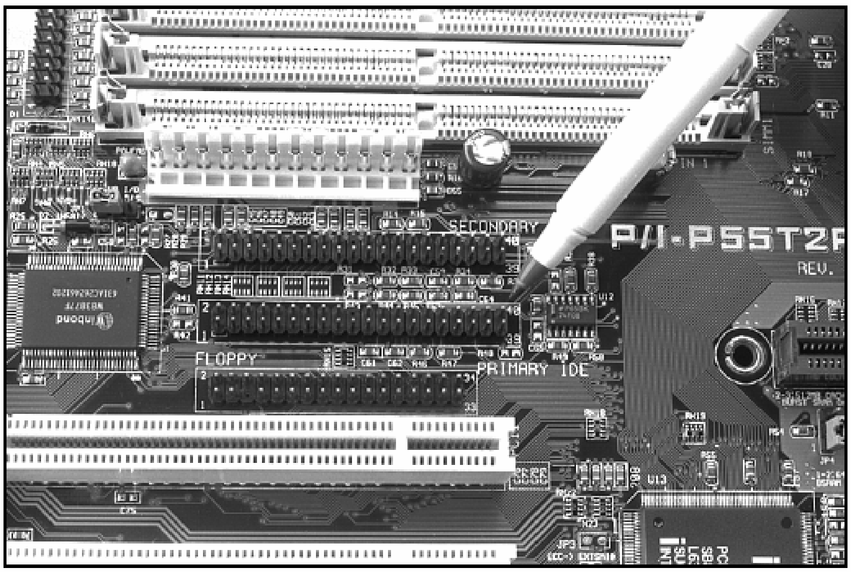


The rear connection interfaces of a typical IDE internal CD-ROM drive.

If the drive is to be the only device on your secondary EIDE interface, the factory settings are usually correct. Consult the manual to make sure this is so.

When you use the CD-ROM or DVD-ROM as a secondary drive—that is, the second drive on the same ribbon cable with another device—make sure it is jumpered as the slave drive, and configure the other device so it is the master drive (see Figure 14.7). In most cases, the drive shows up as the next logical drive, or the D: drive.

Warning
Whenever possible, you should not connect a CD-ROM or DVD-ROM drive to the same channel as a hard disk drive because sharing the channel can slow down both devices. If your computer has the two IDE channels standard in systems today, you should connect the optical drive to the secondary channel, even if you have only one hard disk drive on the primary.



An embedded EIDE interface with primary and secondary IDE connections (the pen is pointing to the primary IDE connector).

Internal Drive Installation

Unpack your internal drive kit. You should have the following pieces:

1. The drive
2. Internal CD-audio cable
3. Floppy disks/CD-ROM with device driver software and manual
4. Drive rails and mounting screws

Troubleshooting Optical Drives

Failure Reading a CD

If your CD fails to read a CD, try the following solutions:

1. Check for scratches on the CD data surface.
2. Check the drive for dust and dirt; use a cleaning CD.
3. Make sure the drive shows up as a working device in System Properties.
4. Try a CD that you know to work.
5. Restart the computer (the magic cure-all).
6. Remove the drive from Device Manager in Windows 9x, allow the system to redetect the drive, and then reinstall the drivers (if PnP-based system).

Failure to Read CD-R, CD-RW Discs in CD-ROM or DVD Drive

If your CD-ROM or DVD drive fails to read CD-R and CD-RW discs, try the following solutions:

1. Check compatibility; some very old 1x CD-ROM drives can't read CD-R media. Replace the drive with a newer, faster, cheaper model.
2. Many early-model DVD drives can't read CD-R, CD-RW media; check compatibility.
3. The CD-ROM drive must be MultiRead compatible to read CD-RW because of the lower reflectivity of the media; replace the drive.
4. If some CD-Rs but not others can be read, check the media color combination to see whether some color combinations work better than others; change the brand of media.
5. Packet-written CD-Rs (from Adaptec DirectCD and backup programs) can't be read on MSDOS/ Windows 3.1 CD-ROM drives because of the limitations of the operating system.

ATAPI CD-ROM Drive Runs Slowly

If your ATAPI CD-ROM drive performs poorly, check the following items:

1. Check the cache size in the Performance tab of the System Properties control panel. Select the quad-speed setting (largest cache size).
2. Check to see whether the CD-ROM drive is set as the slave to your hard disk; move the CDROM to the secondary controller if possible.
3. Your PIO (Programmed I/O) or UDMA mode might not be set correctly for your drive in the BIOS; check the drive specs and use autodetect in BIOS for the best results. Check that you are using busmastering drivers on compatible systems; install the appropriate drivers for the motherboard's chipset and the operating system in use.
4. Check to see whether you are using the CD-ROM interface on your sound card instead of ATA connection on motherboard. Move the drive connection to the

ATA interface on the motherboard and disable the sound card ATA if possible to free up IRQ and I/O port address ranges.

5. Open the System Properties control panel and select the Performance tab to see whether the system is using MS-DOS Compatibility Mode for CD-ROM drive. If all ATA drives are running in this mode, see www.microsoft.com and query on “MS-DOS Compatibility Mode” for a troubleshooter.

If only the CD-ROM drive is in this mode, see whether you’re using CD-ROM drivers in CONFIG.SYS and AUTOEXEC.BAT. Remove the lines containing references to the CD-ROM drivers (don’t actually delete the lines—REM them), reboot the system, and verify that your CDROM drive still works and that it’s running in 32-bit mode. Some older drives require at least the CONFIG.SYS driver to operate.

Trouble Reading CD-RW Discs on CD-ROM

If you can’t read CD-RW discs in your CD-ROM, check the vendor specifications to see whether your drive is Multi-Read compliant. Some drives are not compliant.
If your drive is Multi-Read compliant, try the CD-RW disc on a known-compliant CD-ROM drive (a drive with the Multi-Read feature).

Trouble Reading CD-R Discs on DVD Drive

If your DVD drive can’t read a CD-R disc, check to see that the drive is MultiRead2 compliant—non-compliant DVDs can’t read CD-R media. Newer DVD drives generally support reading CD-R media.

Trouble Making Bootable CDs

If you are having problems creating a bootable CD, try these possible solutions:

1. Check the contents of bootable floppy disk from which you copied the boot image. To access the entire contents of a CD, a bootable disk must contain CD-ROM drivers, AUTOEXEC.BAT, and CONFIG.SYS.
2. Use the ISO 9660 format. Don’t use the Joliet format because it is for long-filename CDs and can’t boot.
3. Check your system’s BIOS for boot compliance and boot order; the CD-ROM should be listed first.

DVD

DVD stands for digital versatile disc and in simplest terms is a high-capacity CD. In fact, every DVDROM drive *is* a CD-ROM drive. It means that they can read CDs as well as

DVDs (discs). DVD uses the same optical technology as CD. The main difference is the higher density. A CD-ROM can hold a maximum of about 700MB (80-minute disc) of data, which might sound like a lot but is simply not enough for many up-and-coming applications, especially where the use of video is concerned. DVD discs, on the other hand, can hold up to 4.7GB (single layer) or 8.5GB (dual layer) on a single side of the disc, which is more than 11 ½ times greater than a CD. Double-sided DVD discs can hold up to twice that amount; although currently you must manually flip the disc over to read the other side.

The 4.7 GB of digital data according to Moving Picture Experts Group–standard 2 (MPEG-2) compression is enough to contain approximately 133 minutes of video, which is enough for a full-length, full-screen, full-motion feature film. Using both layers, a single-sided disc could easily hold 240 minutes of DVD video or more.

DVD technology is similar to CD technology. Both use the same size (120mm diameter, 1.2mm thick, with a 15mm hole in the center).

DVD Drive Compatibility

When DVD drives first appeared on the market, they were touted to be fully backward compatible with CD-ROM drives. Although that might be the case, when reading commercially pressed CD-ROM discs, that was not necessarily true when reading CD-R or CD-RW media. Fortunately, the industry has responded with standards that let you to know in advance how compatible your DVD drive will be. These standards are called *Multi-Read* for computer-based drives and *Multi-Play* for consumer standalone devices, such as DVD-Video or CD-DA players.

Installing a DVD Drive

A DVD drive installs in a manner similar to that of a CD-ROM or any other type of drive.
Following are some basic principles to keep in mind when installing a DVD drive:

1. *Most DVD drives are ATAPI devices.* So, you have the issue of master/slave setups on the dual-connector 40-pin ATA cable. If you're replacing an old CD-ROM drive, simply note its master/slave configuration when you remove it from your system, set the new DVD drive the same way, attach it to the data cable, connect the power, and you're finished.
2. *Many current DVD drives require that you connect them to a bus-mastering ATA interface.*
3. *If you want to play DVD movies using your drive, you need an MPEG-2 decoder.* This can be either a card (hardware decoder) or a program (software decoder). If you are using a hardware decoder, it requires an open PCI expansion slot, as well

as an IRQ (interrupt request) resource from your system. Also, to view the movies on your PC's display, you must somehow connect the decoder card to your existing video card. This can be accomplished via an internal connection to the card, or it can use a loop through connection to the monitor port on the back of the card. Some video cards include MPEG-2 decoders (players) built in.

Note that as processors have become more powerful, many DVD drives now ship with MPEG-2 software decoders. These eliminate the need for an MPEG-2 decoder card, and if your processor is fast enough, it allows for adequate performance when playing DVD video discs. Most of the decoders recommend a minimum of a 400MHz processor, and of course, even faster would be better. Although the software decoders can work well, in most cases if you really plan on using your PC to watch DVD movies, it is worth the extra cost for a hardware MPEG-2 decoder card.

Optical Drive Performance Specifications

Typical performance figures published by manufacturers are the data transfer rate, the access time, the internal cache or buffers (if any), and the interface the drive uses.

Data Transfer Rate

The data transfer rate tells you how quickly the drive can read from the disc and transfer to the host computer. Transfer speed is measured two ways. The one most commonly quoted with CD/DVD drives is the "x" speed, which is defined as a multiple of the particular standard base rate. For example, CD-ROM drives transfer at 153.6KB/sec according to the original standard. Drives that transfer twice that are 2x, 40 times that are 40x, and so on.

DVD drives transfer at 1,385KB/sec at the base rate, whereas drives that are 20 times faster than that are listed as 20x. With recordable CD drives, the speed is reported for various modes. CD-R drives have two speeds listed (one for writing, the other for reading), and CD-RW drives have three. On a CD-RW drive, the speeds are in the form A/B/C, where A is the speed when writing CD-Rs, B is the speed when writing CD-RWs, and C is the speed when reading. The first CD-RW drive on the market was 2/2/6, with versions up to 20/10/40 available today.

Access Time

The access time for a CD or DVD drive is measured the same way as for PC hard disk drives. In other words, the access time is the delay between the drive receiving the command to read and its actual first reading of a bit of data. The time is recorded in

milliseconds; a typical manufacturer's rating would be listed as 95ms. This is an average access rate; the true access rate depends entirely on where the data is located on the disc.

Buffer/Cache

Most CD/DVD drives include internal buffers or caches of memory installed onboard. These buffers are actual memory chips installed on the drive's circuit board that enable it to stage or store data in larger segments before sending it to the PC. A typical buffer for a CD/DVD drive is 128KB, although drives are available that have either more or less (more is usually better). Recordable CD or DVD drives typically have much larger buffers of 2MB–4MB or more to prevent buffer underrun problems and to smooth writing operations. Generally, faster drives come with more buffer memory to handle the higher transfer rates.

A minimum size of 128KB for a built-in buffer or cache is recommended and is standard on many 24x and faster drives. For greater performance, look for drives with 256KB or larger buffers.

CPU Utilization

This "CPU utilization" factor refers to how much attention the CPU must provide to the hardware or software to help it work. A low CPU utilization percentage score is desirable because the less time a CPU spends on any given hardware or software process, the more time it has for other tasks and thus the greater the performance for your system.

Direct Memory Access and Ultra-DMA

Bus-mastering ATA controllers use Direct Memory Access (DMA) or Ultra-DMA transfers to improve performance and reduce CPU utilization. Virtually all modern ATA drives support Ultra-DMA. With bus-mastering, CPU utilization for ATA/ATAPI and SCSI CD-ROM drives is about equal at around 11%. Thus, it's to your benefit to enable DMA access for your CD-ROM drives (and your ATA hard drives, too) if your system permits it. Most recent ATA ATAPI CD-ROM drives (12x and above) support DMA or Ultra-DMA transfers, as does Windows 95B and above and most recent Pentium-class motherboards. To determine whether your Win9x or Me system has this feature enabled, check the System Properties' Device Manager tab and click the + mark next to Hard Disk Controllers. A drive interface capable of handling DMA transfers lists "Bus Master" in the name. Next, check the hard drive and CD-ROM information for your system.

You can use the properties sheet for your system's CD-ROM drives under Windows 9x/Me and Windows 2000/XP to find this information; you might need to open the system to determine your hard drive brand and model. Hard disk drives and CD-ROM

drives that support MultiWord DMA Mode 2 (16.6MB/sec), UltraDMA Mode 2 (33MB/sec), UltraDMA Mode 4 (66MB/sec), or faster can use DMA transfers. Check your product literature or the manufacturer's Web sites for information. To enable DMA transfers if your motherboard and drives support it, open the System Properties sheet in Windows 9x or later, click the Device Manager tab, and open the Properties sheet for your hard drive. Click the Settings tab, and place a check mark in the box labeled DMA.

Repeat the same steps to enable DMA transfers for any additional hard drives and ATAPI CD- ROM drives in your computer. Restart your computer after making these changes.

Advantages of CD-ROM Drives

If a hard disk drive holds more information than a floppy disk drive, accesses the information faster, and reads and writes information, then why do we need CD-ROM drives? The answer is simple: a compact disc can hold large amounts (650 MB) of removable data and can be mass-produced at a very low cost.

The CD has become the medium of choice for software distribution by manufacturers. DVD (digital video disc) technology is beginning to replace traditional CD-ROM technology on many new PCs, but DVD drives can read CD-ROM. It is expected that CD-ROM will be a standard distribution method for the foreseeable future.

An entire software package can be stored on one CD. For example, the early versions of the Microsoft Office Suite were supplied on 32 floppy disks. Today, the entire program suite and its manuals are stored on a single CD. It is also much faster to install a CD. The user simply starts it up, enters any required information, and comes back later; it's no longer necessary to feed disk after disk into the computer. When they were introduced, CDs held large databases such as encyclopedias. Today, they are used for every possible type of data, from national phone directories and software libraries to collections of clip art, music, and games. The following table lists the advantages of storing data on a CD.

Advantage	Description
Large storage capacity	Up to 650 MB of data fit on a single 5-inch disc. (Smaller than the original 5.25-inch floppy disk, a CD holds almost 2000 times as much information.)
Portability	The CD is a portable medium.
Data cannot be changed	A CD is read-only, which prevents accidental erasure of programs or files.
Sturdiness	More durable than the standard 5.25-inch or 3.5-inch disks, CDs

	are not magnetic media and thus are not subject to the same dangers posed by proximity to electrical sources or magnets.
Special capabilities	CD-ROMs are audio-capable, allowing special compression of audio, image, and video data. They can be used to play standard audio CDs and have the capacity to store and record video data.

About CD-ROM Standards

The CD-ROM world makes use of several standards. These are usually referred to by the color of the cover of the volume issued by the ISO (International Organization for Standardization) committee—for example, the White Book, Yellow Book, and so on. ISO formats are discussed in more detail later in this lesson.

CD-ROM Technology

CD-ROMs store data as a series of 1s and 0s, just like a floppy disk or a hard disk drive. However, instead of using magnetic energy to read and write data, CD readers and writers use laser energy. There are two major advantages to using lasers:

- There is no physical contact between the surface of the CD and the reading device.
- The diameter of the laser beam is so small that storage tracks can be written very close together, allowing more data to be stored in a smaller space.

Connecting a CD-ROM Drive

A CD-ROM drive is a peripheral device and must be connected to the bus of the computer through a controller. There are several ways to install a CD-ROM drive.

IDE

New computers have primary and secondary IDE connectors as part of the motherboard and BIOS setup. It is becoming commonplace to install CD-ROM drives on the secondary controller.

Multimedia

The term *multimedia* embraces a number of computer technologies, but refers primarily to video, sound, and the storage required by these large files. Basically, multimedia is a combination of graphics, data, and sound on a computer. In all practicality, the concept of

adding multimedia simply means adding and configuring a sound card, a video card, and a CD-ROM drive to a system.

Microsoft formed an organization called the Multimedia PC Marketing Council in 1991 to generate standards for multimedia computers. The council created several multimedia PC (MPC) standards, and it licenses its logo and trademark to manufacturers whose hardware and software conform to these guidelines.

The Multimedia PC Marketing Council formally transferred responsibility for its standards to the Multimedia PC Working Group of the Software Publishers Association (SPA). This group includes many of the same members as the original MPC Marketing Council. The group's first creation was a new MPC standard.

The MPC Marketing Council originally developed two primary standards for multimedia: MPC Level 1 and MPC Level 2. Under the direction of the SPA, the first two standards have been replaced by a third, called MPC Level 3 (MPC 3), which SPA introduced in June 1995. (There are currently no plans for the publication of any additional MPC standards.) These standards define the minimum capabilities for a multimedia computer. The following table presents these standards.

	MPC Level 1	MPC Level 2	MPC Level 3
Processor	16 MHz 386SX	25 MHz 486SX	75+ MHz
RAM	2 MB	4 MB	8 MB
Hard disk	30 MB	160 MB	540 MB
Floppy disk	1.44 MB 3.5-inch	1.44 MB 3.5-inch	1.44 MB 3.5-inch
CD-ROM	Single-speed	Double-speed	Quad-speed
Audio	8-bit	16-bit	16-bit
VGA video resolution	640 x 480; 16 colors	640 x 480; 64,000 colors	640 x 480; 64,000 colors
Other I/O	Serial; parallel; MIDI; game	Serial; parallel; MIDI; game	Serial; parallel; MIDI; game
Software	Microsoft Windows 3.1	Microsoft Windows 3.1	Microsoft Windows 3.1
Date introduced	May 1993	1994	June 1995

You should consider the MPC 3 specification as the bare minimum for any multimedia system today. Specifically, a recommended system exceeds the Level 3 standards in several areas such as RAM, hard disk size, and video capability. Note that although speakers are not technically part of the MPC specification, sound reproduction does require external speakers! The built-in speaker used for POST beep codes is not sufficient for this quality of sound.

Video-Capture Software

With the advent of multimedia computers and software, manipulating full-motion video was the next logical step. A modern high-speed multimedia computer has become standard equipment in the moviemaking industry. Today, even amateur filmmakers can use their computers to give home movies a touch of professionalism.

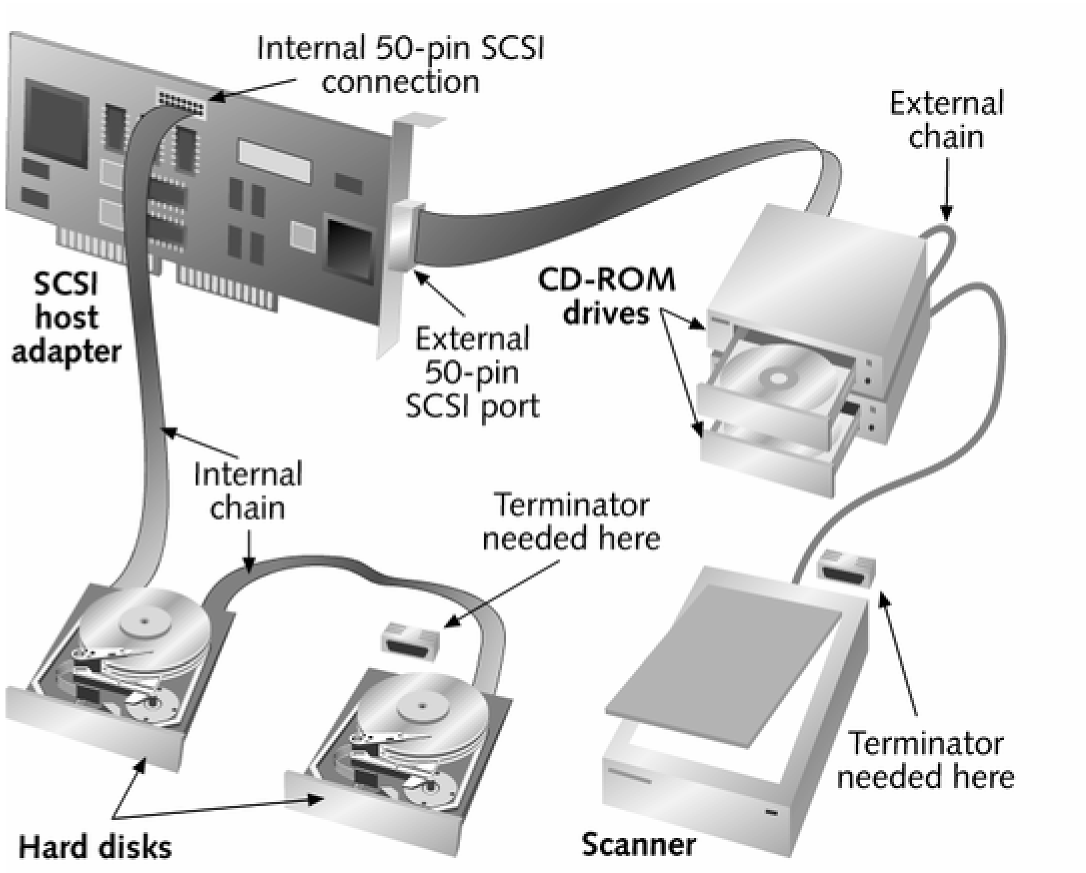
Video-capture software provides an interface that allows users to import and export video formats in order to edit them with their computers. This software allows a user to view audio waveforms and video images, create files, capture single frame or full-motion video, and edit video clips and still frames for content and effects.

File-editing functions such as zoom, undo, cut, paste, crop, and clear can be used to edit audio and visual files. Users can also set the compression controls to the type of format desired and determine the capture rates. The capture rate for full-motion video (equivalent to what you would find in TV or on the big screen) is 30 frames per second (fps), but some systems might not be able to reach this potential. Professional systems include very large, very fast hard disk drives for data buffering. A typical user of video-capture software might realize a frame-capture rate of only up to 15 fps without adding an arsenal of hardware to enhance the system.

NOTE

Most new PCs will far exceed the basic multimedia requirements listed above. The A+ test current at the time of writing should not, however, go beyond these features. For now, you should still be able to describe the MPC features for the exam.

SCSI Drives



The Small Computer System Interface (SCSI) has become the mass-storage device of choice for large network installations. SCSI has many advantages over standard IDE and EIDE drives. SCSI is the favored drive for high-end workstations, network and Internet servers, and the Macintosh line of personal computers. In many installations, the advantages far outweigh the slight extra effort in configuration. In this lesson, we explore the advantages and uses of a SCSI system.

SCSI was introduced in 1979 as a high-performance interface, allowing connection of both internal and external devices. Because it runs on virtually any operating system, it was adopted by the ANSI Standards Committee and is now an open standard in its third generation.

At its core, SCSI is a simple design. A single card, the host adapter, (or a chip set on the motherboard) connects up to 15 devices. These devices can be attached inside or outside the PC using standard cables and connectors. SCSI is the only interface that can connect such a wide variety of devices. Communication between the devices and the host adapter

is done without involving the CPU or the system bus until data must be passed to one or the other.

This design frees expansion slots and reduces the number of interrupts and memory addresses needed, while cutting down the number of drivers required. Less-robust solutions, such as IDE and EIDE, are little more than switching stations, relying on the PC's CPU to manage the data bus. SCSI host adapters are true subsystems with advanced commands that can order and route data to improve performance.

SCSI-3

To speed up the pace of development, the SCSI Committee approved a "fast track" system for the SCSI-3 standard. It let a subcommittee handle most of the work, and new subsections were adopted without waiting for the publication of the entire SCSI definition.

That bright idea, plus the advent of the PCI bus and mature Plug and Play operating systems, has made it easy to install components and given users excellent control and flexibility. All SCSI-3 cards have ways to support existing SCSI-2 devices. Some of the highlights of current state-of-the-art SCSI technology on the desktop include the following seven features:

High-Performance Products

The success and stability of the SCSI standard makes it an ideal platform for developing high-performance products. SCSI's robust, reliable interface and advanced commands allow manufactures to build "best-of-breed" products to take advantage of its power. The fastest hard disk drives and CD-ROM devices traditionally show up first, sporting a SCSI interface. The most advanced scanners are SCSI-based, and many optical products come only in SCSI versions. Even when non-SCSI versions reach the market, they generally under-perform their SCSI siblings.

Plug and Play Installation

Well-designed SCSI cards are recognized and drivers are installed automatically with Plug and Play operating systems such as Windows 98, Windows NT, 2000, and the Macintosh OS. Most SCSI-based peripherals provide Plug and Play setup. The first time the system is booted up after they are added to a SCSI chain, the system notices the new device and asks for the product's setup disk.

Simple Expansion

Adding external devices is as simple as connecting an industry standard cable and power cord. If users decide to add additional host adapters, they can share the same drivers, reducing system overhead.

Advanced Management

SCSI products generally offer a range of tools to tune the bus and devices attached to it. For example, many host adapters have firmware that provides the ability to format and inspect hard disk drive reliability and define custom settings for each device on the chain. Operating-system utilities are provided to check the status of a device and enable advanced features.

SCAM Support

This SCSI acronym stands for "SCSI configured auto-magically." Most new SCSI products are SCAM enabled, meaning that the user does not have to worry about setting the ID numbers for them, because they will configure themselves, using an open ID position on the SCSI chain.

NOTE

Even if a hard disk drive is SCAM enabled, you might have to set an ID on multidrive PCs, because the host adapter will need it to determine which drive is the normal boot device.

Connect/Disconnect

This command allows a SCSI device handling a large amount of data or performing complex operations to disengage from the host adapter's bus while performing the task, allowing other devices free access until it is finished.

Tag Command Queuing

SCSI devices with this feature can reorder how blocks of data are moved on the bus to speed transfer. The way it functions can be compared to letting a shopper with only a few items move to the head of the checkout line, to reduce the average wait time per shopper.

SCSI and IDE Compared		
Feature	**SCSI**	**UDMA/IDE**
Devices per channel	7/15 per chain	2 per chain

Maximum potential throughput for major classes of SCSI and IDE	160 MB per second (Ultra 160) 80 MB per second (Ultra2) 40 MB per second (Wide SCSI)	66 MB per second (UDMA) 33 MB per second (UDMA) 16.7 MB per second(Fast ATA)
Connection types	Internal and external	Internal only
True bus mastering	Yes	No
Operate more than one I/O device at a time?	Yes	No
Advanced commands (such as tag command queuing, connect/disconnect)	Yes	No

Noise and SCSI

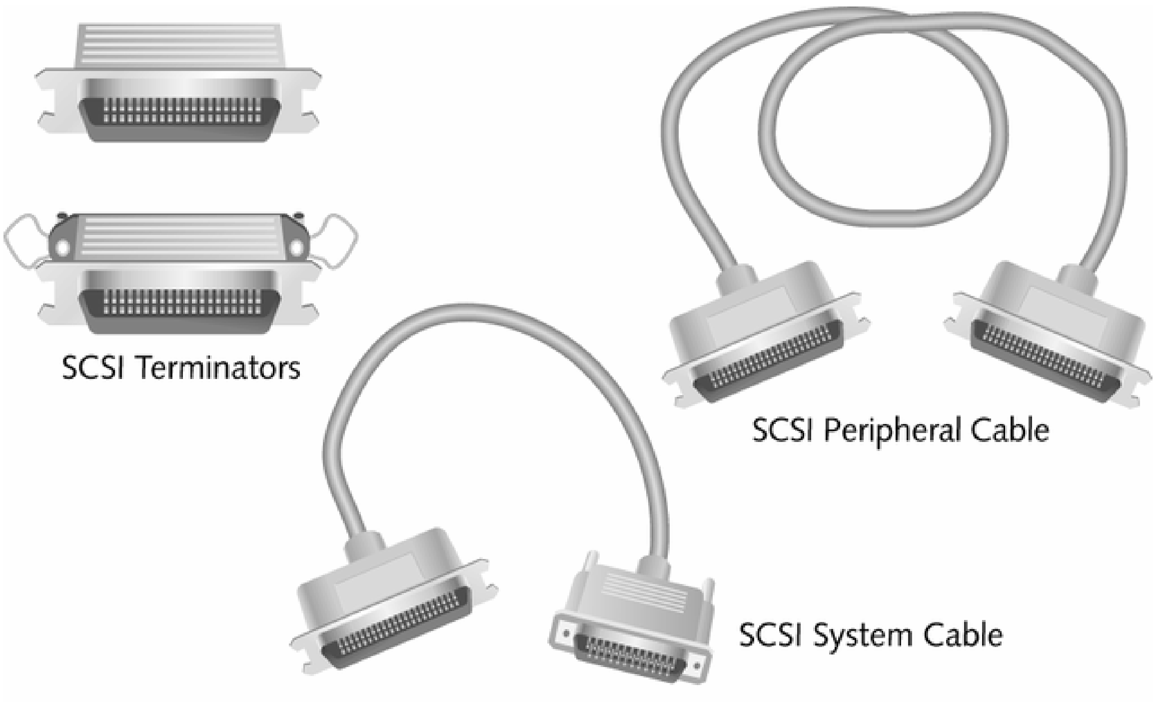
Any electrical signal other than data is called *noise*. Due to the many signals and electrical devices present, the interior of a computer is a noisy place. Computer manufacturers do many things to contain the noise inside the case, including adding shielding and grounding. Anything inside, or directly connected to, a computer is either a contributor to or a victim of the noise.

Because of the high data transfer speed, products using the SCSI-2 and later standards can be very sensitive to noise. Cables tend to act as antennae for noise. For this reason, proper cabling and minimizing of cable length are needed to maintain low noise in a SCSI system. Any noise spread through either the electrical power cables or the data cable is called "common-mode" noise.

A single-ended device communicates through only one wire per bit of information. This one wire is measured, or referenced, against the common ground provided by the metal chassis. Single-ended devices are vulnerable to common-mode noise (they have no way of telling the difference between valid data and noise). SCSI-1 devices are all single-ended.

Some SCSI-2 and SCSI-3 devices are differential-ended. These products employ two wires per bit of data—one wire for the data and one for the inverse of the data. The inverse signal takes the place of the ground wire in the single-ended cable. By taking the

difference of the two signals, the device is able to reject common-mode noise in the data stream.



Troubleshooting a Device Conflict

Determine which is the offending device by taking the following measures:

- Load only the device drivers for the SCSI devices.
- If the problem still occurs, use the F8 key to determine which driver conflicts. (Press F8 when starting MS-DOS or Windows 95 or 98—this will allow step-by-step confirmation of the startup process.)
- If the device driver is an executable file, try running it with the "/?" option. This will usually show a variety of command-line switches for the device driver

Here are some ways to correct the problem after you've found it:

- Look in the manuals or "readme" files of both devices. The problem might be a common one with a known solution.
- Try a variety of switches to see if any of them solves the problem.
- Attempt to find an updated driver for one or both of the devices (the Internet is a good place to look).
- If none of those solutions fixes the problem, you might be forced to choose between the devices or go to a multiple boot configuration.

Memory Management

SCSI host adapters typically have their own ROM chips. For MS-DOS systems, put the appropriate "X=" statements in the EMM386.EXE line of the CONFIG.SYS and the appropriate EMMEXCLUDE= statement in the SYSTEM.INI file. (For more details about configuring these files, A missing or erroneous "exclude" statement can cause intermittent lock-up problems.

Setting Up a SCSI Subsystem

There are several steps in setting up a SCSI-based system or adding a new SCSI peripheral to an existing system. Performing these steps in the proper order, without shortcuts, is the key to a fast, easy installation.

Start with the Host Adapter

SCSI cards come in a wide variety of sizes, shapes, and configurations. Some offer one connection, others have four. Options include secondary or even tertiary channels—RAID, cache RAM, and so forth. Be sure that the card will be able to service the devices planned for it. Set any jumpers first, then install the SCSI adapter card in the appropriate expansion slot.

Set the SCSI IDs, Termination, and Peripheral Cabling

Write down the ID for each device—including the host adapter—as it is assigned. After the IDs are set, verify termination for each end of the chain. Finally, attach the cables—first to the host adapter, then to the closest internal device—and move outward on the chain. Repeat the process for the external devices.

External devices usually use some form of switch to set the ID. Most allow setting IDs from 0 through 7 only. You might need to adjust that with internal devices that often allow a wider range of ID numbers. Cable types include: 50-pin Centronics type, SCSI-2 D-Shell 50, and 68-p type connectors. Make sure the last device in the chain is properly terminated.

Internal SCSI devices are installed inside the computer and are connected to the host adapter through an interior connector on the host adapter. Check the connection diagram to be sure the fitting is the right one for that type of device. The options are a 50-pin ribbon cable (similar to a 40-pin IDE cable) and two similar 68-pin cables. Be sure to use the right type of 68-pin cable: one is for ultra-low voltage differential (LVD), and the other is for single-ended (SE) drives. They are NOT interchangeable.

IMPORTANT

SCSI devices connected incorrectly (for instance, with the cable plugged backwards) can be damaged! Be sure the red or blue strip on the cable is facing toward pin 1. Some SCSI devices allow only a proper connection.

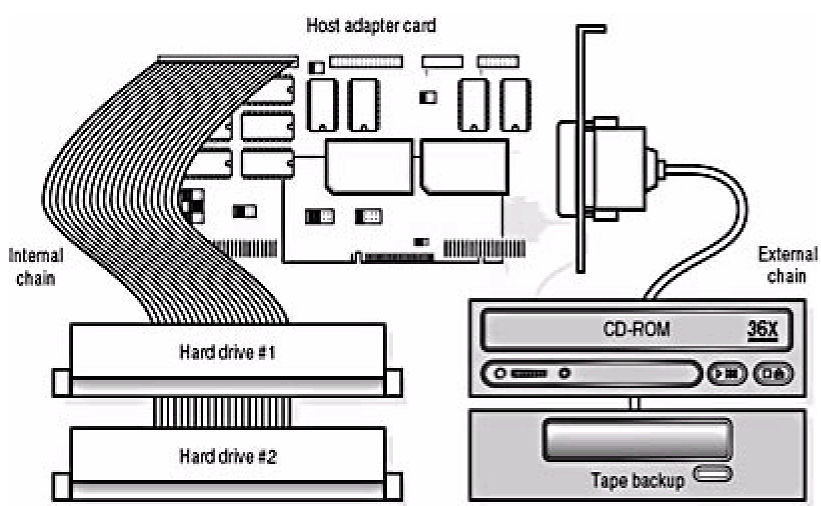
Power Up One Device at a Time

A good practice is to connect the power to one device, power up, and check for problems. Then power additional devices one at a time and make sure everything is working and without conflict.

Load Operating System Drivers and SCSI Software

Finally, load any software required to allow the operating system to recognize the new hardware and take full advantage of its features.

Using a cable with enough connectors enables you to easily link multiple internal devices. You can have up to eight (numbered 0 through 7) devices, or 16 (numbered 0 through 15, depending on the host adapter and the devices) on a single SCSI chain. Don't forget that one position in each SCSI chain is taken up by the host adapter. The picture below shows a SCSI chain.



SCSI Chain

The exact number of devices will vary depending on a number of conditions. The host adapter must support the number selected, the installer must be able to set proper IDs, and the cables and connectors must be compatible. Older adapters allow only seven total IDs, and the card will use one, leaving you with six devices. Some SCSI devices have limited ID options. Many older products have only seven possible settings; some

scanners or optical products are factory-set to an ID. Given the range of cable options and performance considerations, you might have to limit the number of devices on a single chain to get maximum performance.

Setting SCSI IDs

A simple SCSI chain works like a network, and—like a network—each device requires its own unique address. But unlike a network, setting an address on a SCSI chain is simple. A SCSI device can have any ID number in a range recognized by the host adapter, as long as no other device on the same chain has been set to the same number.

In SCSI numbering conventions:

- The host adapter is typically set to SCSI ID 7. (This is a de facto standard, not a requirement.)
- There is no mandated order for the use of SCSI IDs, but the SCAM feature will use a preestablished pattern of IDs if one is available.
- The ID of a bootable hard disk drive may be preset by the host adapter manufacturer. Most manufacturers use SCSI ID 0, although a few are configured to SCSI ID 6.

Setting a SCSI ID for a device is accomplished using jumpers or switches located on, or inside, the SCSI device. Typically, all internal SCSI hard disk drives use jumpers to set their IDs. External devices usually (but not always) have switches. Some SCSI devices have automatic ID and termination, using SCAM.

If you plan to utilize a SCSI drive as your C drive (this is required if you want to boot into MS-DOS from this drive), it must be configured as a bootable drive. This is done by either specifying the host adapter as the "bootable" SCSI ID or setting the host adapter to emulate a standard AT-style controller.

Logical Unit Numbers

It is possible to have a single SCSI ID support more than one device. *Logical unit numbers* can be used to provide a unique identifier for up to seven subunits per ID number. These are used primarily in hard disk drive arrays to create one large logical drive out of several smaller physical drives. Logical Unit Numbers require highly specialized software and are most often found in network servers running NetWare, Windows NT, or UNIX.

Chapter 10

Expansion Buses

OBJECTIVES

Understanding Expansion Buses

Configuring Expansion Cards

Understanding Expansion Buses

Expansion cards are the small circuit boards that you put into slots on the motherboard that make your computer do "neat" things. Such cards are video cards, sound cards, modems, image capture cards, etc. Cards are one of the most simple topics in computer hardware. You buy the kind of card you want, stick it in the slot, and load the new drivers, Buses including modems, video cards, and portable drives. We also discuss conflicts within the computer how they are created and reconciled.

Expansion buses are used to connect devices to the motherboard using the motherboard's data bus. They allow the flow of data between that device and other devices in and connected to the computer. Early computers moved data between devices and the processor at about the same rate as the processor. As processor speeds increased, the movement of data through the bus became a bottleneck. Therefore, the design capability of the buses needed to evolve, too. This lesson discusses that evolution.

Development of the Expansion Bus

As discussed earlier "The Central Processing Unit," every device in the computer RAM, the keyboard, network card (NIC), sound card, and so forth is connected to the external data bus. Expansion slots on the motherboard are standardized connections that allow the installation of devices not soldered to the motherboard. The function of an expansion slot is to provide configuration flexibility when devices are added to a computer.

Whether a device is soldered to the motherboard or connected through an expansion slot, all integrated circuits are regulated by a quartz crystal. The crystal sets the timing for the system, giving all parts access to a common reference point for performing actions. Most CPUs divide the crystal speed by two. (If the CPU has a 33-MHz speed, a 66-MHz crystal is required.) Every device soldered to the motherboard—keyboard chip, memory controller chip, and so on—is designed to run at the speed (or at half the speed) of the system crystal.

CPU speeds increased as technology improved, while the speeds of expansion cards remained relatively constant. It was not practical to redesign and replace every expansion card each time a new processor was released—this would have been complicated and expensive for manufacturers. (And, of course, the additional expense would have been passed along to the consumer.) The commitment of the industry to maintain backward compatibility further complicated design tasks, because any new technology would have to run the older, slower devices.

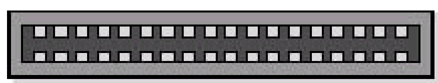
To resolve this dilemma, designers have divided the external data bus into two parts:

- **System bus**: This supports the CPU, RAM, and other motherboard components. The system bus runs at speeds that support the CPU.
- **Expansion bus**: This supports any add-on devices by means of the expansion slots and runs at a steady rate, based on the specific bus design.

Dividing the bus enhances overall system efficiency. Because the CPU runs off the system clock, upgrading a CPU requires changing only the timing of the system bus, while the existing expansion cards continue to run as before. There is usually a jumper setting that changes the system clock speed to match the CPU. The ability of the motherboard to make this change sets the limit for the processor speed. Next, we take a look at the evolving types of expansion buses.

Industry Standard Architecture (ISA)

The first-generation IBM XT (with the 8088 processor) had an 8-bit external data bus and ran at a speed of 4.77 MHz. These machines were sold with an 8-bit expansion bus (PC bus) that ran at 8.33 MHz.



8-bit PC bus slot

IBM took steps that fueled the rapid development of the personal-computer market. Their engineers framed the PC's design as an open system, using standard, off-the-shelf components. That allowed third-party developers to manufacture cards that could snap into the PC bus. IBM also allowed its competitors to copy the PC bus.

With this move, IBM established the *Industry Standard Architecture (ISA)* interface, thus generating the market for clones. A host of third-party developers worked to create products that enhanced the basic machine's features and to keep prices much lower for add-ons than competing proprietary systems such as those from Apple. Without this push, the PC market would have grown more slowly and probably would have been limited to businesses with the money to pay for the more expensive products.

IBM wanted to include a expansion bus—one that would be compatible with previously released devices—with the release of its AT (Advanced Technology) PC, featuring Intel's 80286 16-bit processor. To accomplish this, the designers added a bus that allowed insertion of either an 8-bit card or a 16-bit card. This change resulted in the standard 16-bit ISA slot. This 16-bit bus officially ran at a top speed of 8.33 MHz, but on some Peripheral Component Interconnect (PCI)-based systems the actual rate for ISA slots proved to be as high as about 10 MHz. (PCI is discussed later in this lesson.)

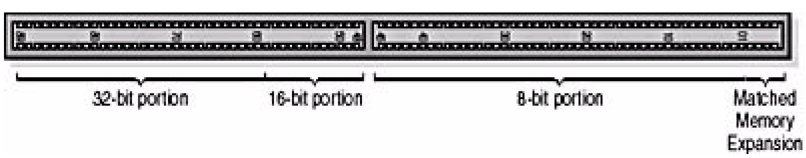
Micro Channel Architecture (MCA)

In 1986, the market came to be dominated by the new 386 machines with their 32-bit architecture. Most PC manufacturers stuck to the same basic ISA design and MS-DOS. Expansion devices based on ISA technology for the 286 AT class machines could be placed in a new 386 clone without problems.

IBM designers produced a new version of the PC, the PS/2 (Personal System/2) and created a proprietary expansion bus called *Micro Channel Architecture (MCA)* as part of the design. Running at 10 MHz, it offered more performance and provided a 32-bit data path. It was also totally incompatible with older ISA cards.

A feature of MCA is its ability to "self-configure" devices. Unlike devices that use technology in which the PC configures itself automatically to work with peripherals such as monitors, modems, and printers, an MCA device always comes with a configuration disk. When installing a new device in an MCA computer, insert the configuration disk (when prompted), and the IRQs, I/O addresses, and DMA channels will be configured automatically. (IRQs, I/O addresses, and DMA channels are discussed in detail in the next lesson.) An MCA bus is shown in the picture below.

The PS/2 never gained enough market share to compete with the 386. MCA cards were few and far between, and more expensive than competing interface designs.



MCA bus

MCA is now a lost technology. As a computer technician, you will not encounter MCA on new computers. However, it is still found in some older machines, and you will need to know how to identify it. If a customer brings in a PS/2 machine for service, be sure to obtain the configuration disks for the computer as well as any MCA cards that go with it.

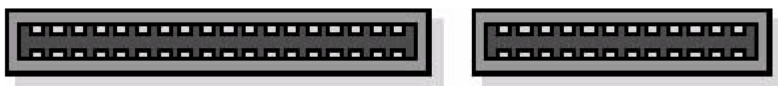
Enhanced ISA (EISA)

In 1988, an industry group answered the challenge of MCA and released a new open standard called *Enhanced ISA* (EISA—pronounced "ee-suh"). It's a 32-bit, 8-MHz standard.

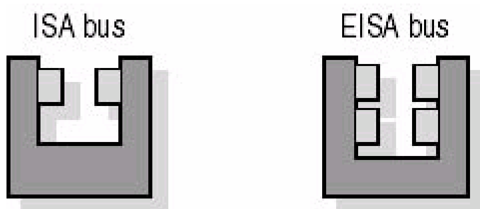
Unlike MCA, EISA uses a variation of the ISA slot that accepts older ISA cards, with a two-step design that uses a shallow set of pins to attach to ISA cards and a deeper

connection for attaching to EISA cards. In other words, ISA cards slip part-way down into the socket; EISA cards seat farther down.

Be very careful to line up cards being placed in an EISA slot precisely and push straight down! If you try to angle the card in, it can be very difficult to seat and you might damage either the connector or the slot. Although EISA is faster and cheaper than MCA, it never gained much more acceptance than MCA. Confusion between MCA and EISA technology—along with a limited need for cards that ran at the faster rate and the fact that only a few display, drive controller, and network cards were made available—led to the early demise of both bus technologies. The pictures below shows how the slot design of the two technologies differs.



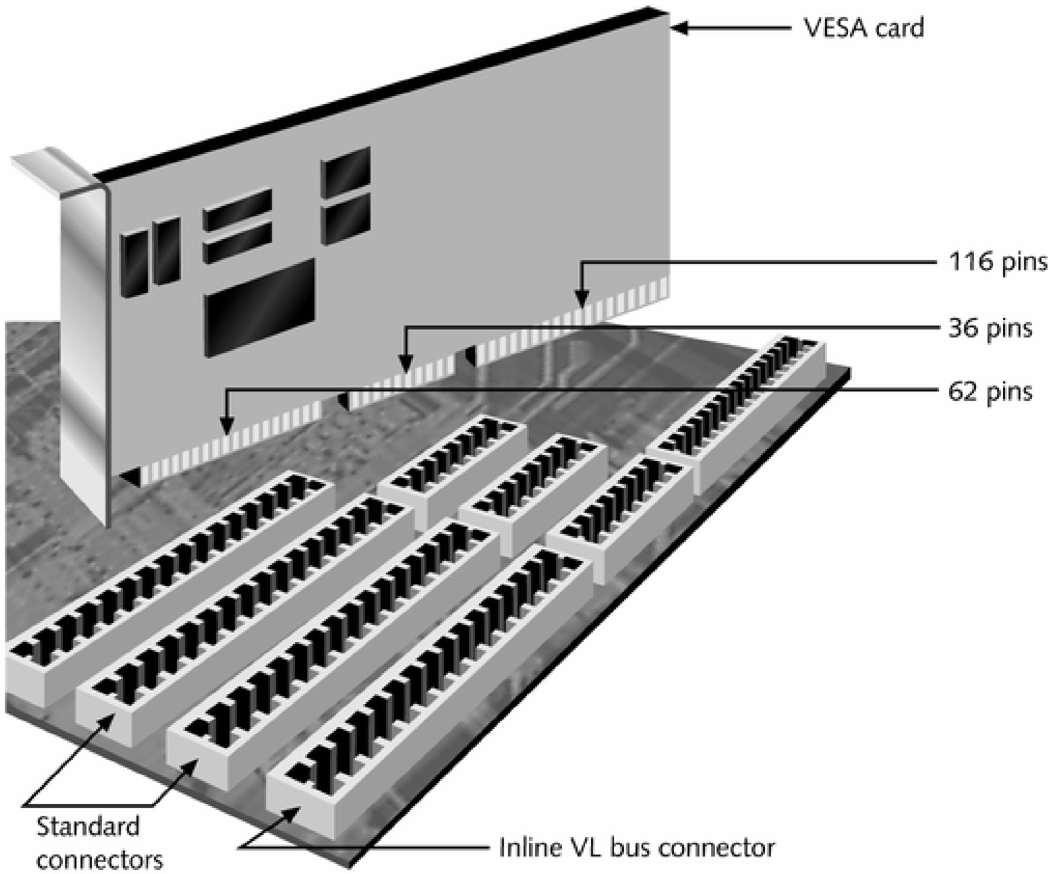
Top view of ISA and EISA bus



Cross section of ISA and EISA bus

VESA Local Bus (VLB)

The problems posed by MCA and EISA designs meant that developers needed an improved bus architecture to speed up graphics adapter performance and keep up with the evolving technologies as new 32-bit operating systems such as Microsoft Windows 3.1 gained popularity. The latter's graphical user interface (GUI) required a much faster display adapter, because every pixel (not just lines of character data) had to be represented and refreshed. About the same time, laser printers and graphics programs like PageMaker and CorelDRAW fostered the "desktop publishing revolution." The hardware industry developed the *VESA local bus (VLB)* to meet the need for a faster expansion interface. (VESA, the Video Electronics Standards Association, was the driving force behind the new bus technology.) Found only in 386 and 486 machines, the VLB had a short lifespan. The cards based on this design are connected directly to the system-bus side of the PC's external data bus.



VESA local bus design

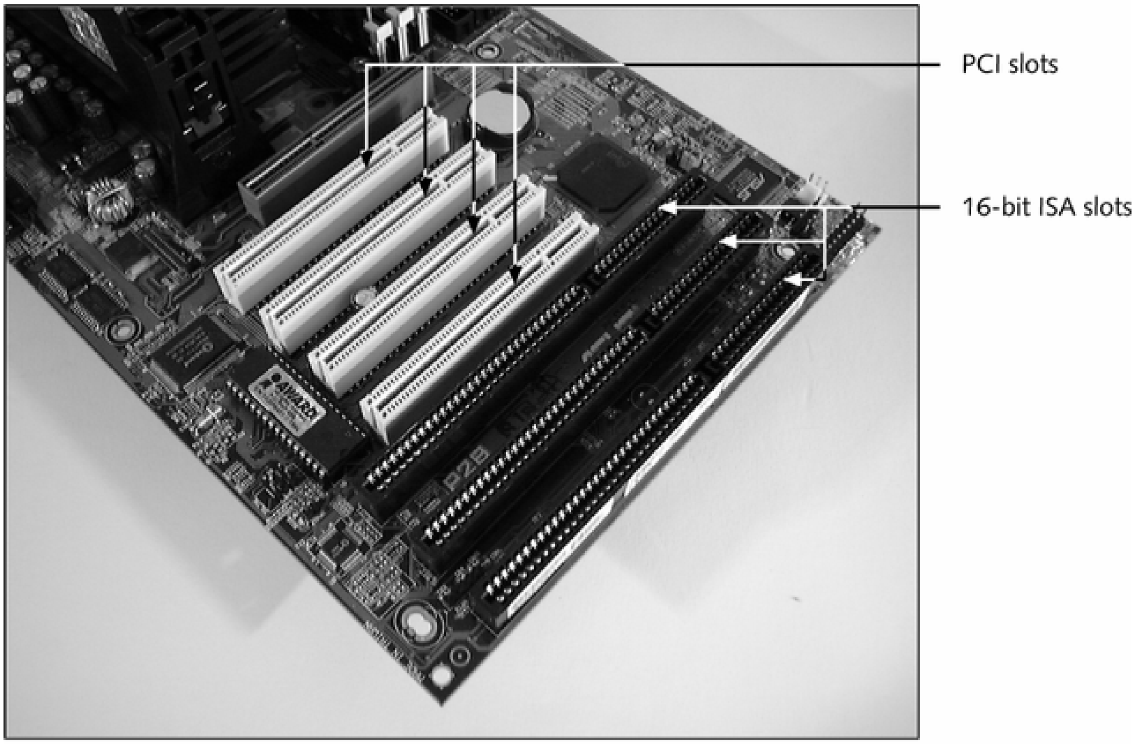
The speed of the system data bus is based on the clock rate of the motherboard's crystal. During the heyday of the VLB, this was usually 33 MHz, and VLB cards usually ran at half that rate, far outpacing the ISA bus. Some cards ran as fast as 50 MHz, using the full speed of the souped-up system bus. That often caused system crashes, because 50 MHz was outside the VLB specification.

The chip design for the VLB controller was relativity simple, because many of the core instructions were hosted by the ISA circuits already on the motherboard, but the actual data passes were on the same local bus as the one used by the CPU.

The design specification provides two other performance-boosting features: *burst mode* and *bus mastering*. In burst mode, VLB devices gain complete control of the external data bus for up to four bus cycles, passing up to 16 bytes (128 bits) of data in a single burst. The actual connectors on the motherboard resemble an ISA slot with an additional short slot aligned with it. On systems that support this interface, one to three slots are located on the side of the motherboard closest to the keyboard connection.

Peripheral Component Interconnect (PCI)

Peripheral Component Interconnect (PCI) allows developers to design cards that will work in any PCI-compatible machine. It overcomes the limitations of ISA, EISA, MCA, and VLB, and it offers the performance needed for today's fast systems.



At first glance, there are many similarities between PCI and the older VLB specifications. Both are local bus systems with 32-bit data paths and burst modes. Also, the original PCI design operates at 33 MHz—roughly the same speed as the VLB. But the important differences between them gained PCI its dominant role in expansion-bus technology. These differences stem from the following features:

- The PCI design's special bus and chip set are designed for advanced bus mastering techniques and full arbitration of the PCI local bus. This allows support of more than three slots.
- The PCI bus has its own set of four interrupts, which are mapped to regular IRQs on the system. If a PC has more than four PCI slots, some will be sharing interrupts and IRQs.

Most PCs on the market today have one or more ISA slot for backward compatibility; however, most expansion cards are now built using the PCI interface. Although Intel was the original driving force behind PCI development, a PCI standards committee maintains

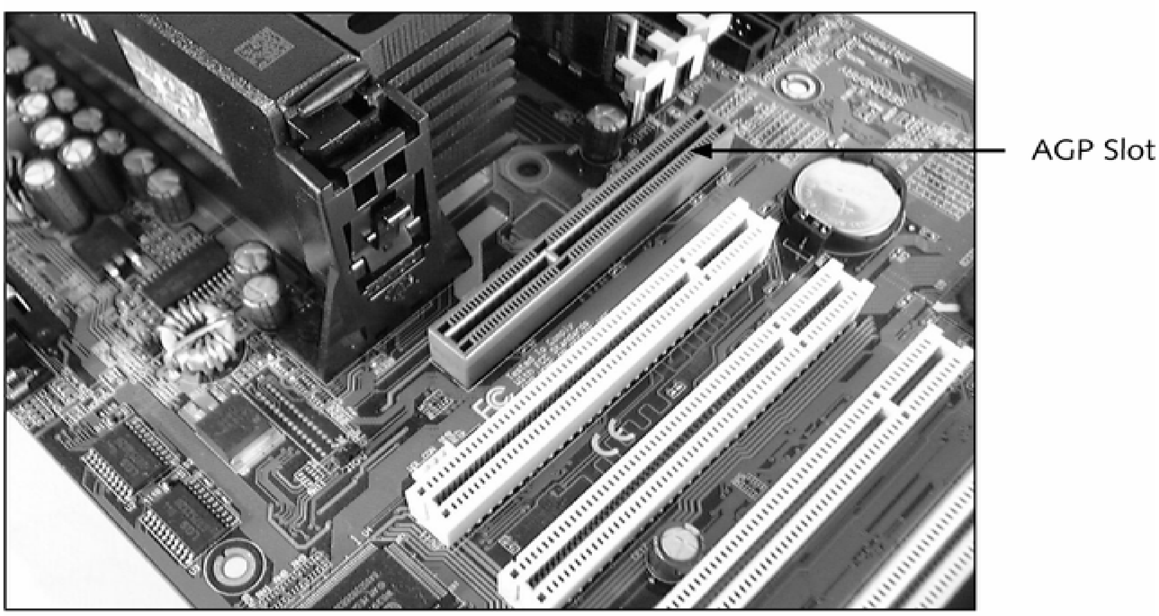
the specification, and it is an open design—anyone can design hardware using PCI without being required to pay royalties.

PCI has gone through many changes, and there are some features to be aware of when you work with one:

- The early PCI motherboards often have jumpers and BIOS settings that must be set to enable proper PCI operation. These are most often found on Pentium 60-MHz and 66-MHz machines.
- The PCI bus speed is not fixed. Newer chip sets can drive it—and the cards on it—at 66 MHz. At full performance, the PCI bus can deliver data transfers at up to 132 MB per second.

Accelerated Graphics Port (AGP)

In the early days of PCI, the major market for that technology was the high-performance display adapter. The popularity of PCI led to its dominance of the expansion-bus market for card manufacturers. Today, the PCI market includes network cards, sound cards, SCSI adapters, UDMA controllers, and DVD interfaces. That posed a problem for display-card designers: having more cards on a single bus slowed down the performance, just when the increasing popularity of 24-bit graphics and 3D rendering called for greater demands on the display system. The search was on for yet another interface; this time, the solution was a single slot—tuned for the display adapter. Once again, Intel led the way and developed the *Accelerated Graphics Port (AGP).*



An AGP slot on the motherboard

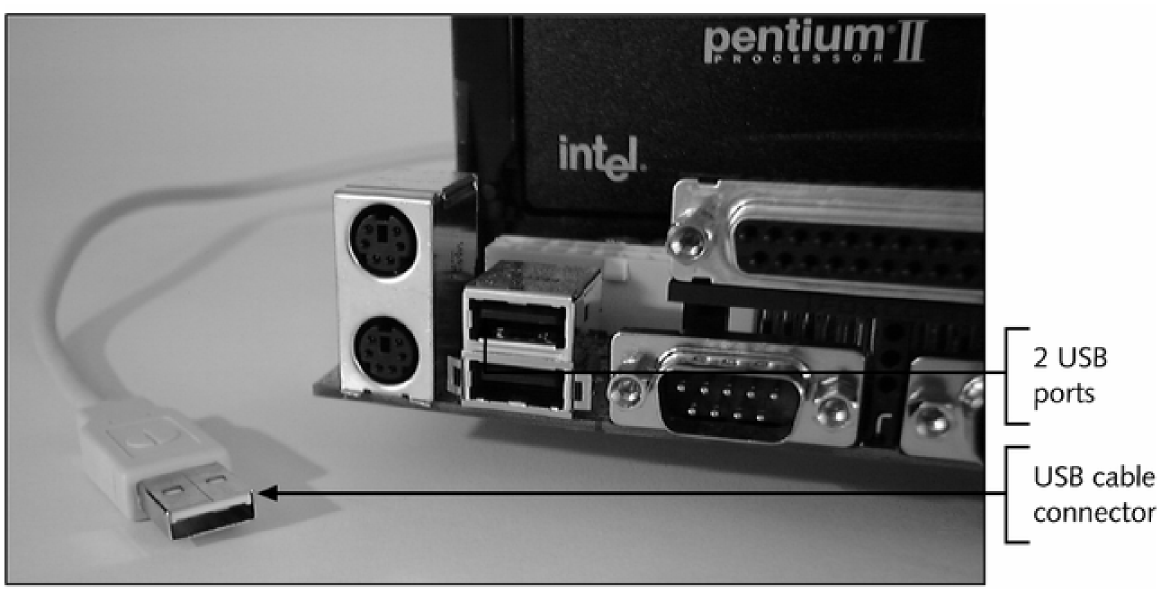
The AGP slot, if present, is the only one of its kind on the motherboard and is usually the slot closest to the keyboard connector. It is set forward of the back PC's case than the PCI slots. APG connectors are found only on Pentium II-based and later computers or on similar CPUs from non-Intel vendors.

Personal Computer Memory Card International Association (PCMCIA)

This is a special socket in which you can plug removable credit-card size devices. These circuit cards can contain extra memory, hard drives, modems, network adapters, sound cards, etc. Mostly, PCMCIA cards are used for laptops, but many PC vendors have added PCMCIA sockets to their desktop machines. The socket uses a 68 pin interface to connect to the motherboard or to the system's expansion bus.

There are three types of PC cards: Type 1 slots are 3.3mm thick and hold items such as RAM and flash memory. Type 1 slots are most often seen in palmtop machines or other handheld devices. Type 2 is 5mm thick and I/O capable. These are used for I/O devices such as modems and network adapters. Type 3 is 10.5mm thick and used mainly for add-on hard drives. When buying PC Card equipment, you must consider the size of the slot. In most cases, Type 3 can handle Type 2 and Type 1.

Universal Serial Bus (USB)



The newest addition to the PC bus collection, the *universal serial bus (USB)* connects external peripherals such as mouse devices, keyboards, joysticks, scanners, and digital cameras to the computer. The USB port is a thin slot; most new motherboards offer two, located near the keyboard. They can also be provided through an expansion card.

USB supports *isochronous* (time-dependent) and *asynchronous* (intermittent) data transfers. Isochronous connections transfer data at a guaranteed fixed rate of delivery. This is required for more demanding multimedia applications and devices. Asynchronous data can be transferred whenever there is no isochronous traffic on the bus. USB supports the following data transfer rates, depending on the amount of bus bandwidth a peripheral device requires:

- 1.5 megabits per second (Mbps) for devices, such as a mouse or keyboard, that do not require a large amount of bandwidth.
- 12 Mbps isochronous transfer rate for high-bandwidth devices such as modems, speakers, scanners, and monitors. The guaranteed data-delivery rate provided by isochronous data transfer is required to support the demand of multimedia applications and devices.

USB devices can be attached with the computer running. A new device will usually be recognized by the operating system, and the user will be prompted for drivers, if drivers are required. Bear in mind that USB is a new standard, and some early USB ports and chip sets do not properly support some newer devices. Problems with embedded USB ports are not generally worth repairing. It is usually better to install a new USB interface card.

Configuring Expansion Cards

In the previous lesson, we discussed the different kinds of expansion buses. The purpose of these buses is to accept expansion cards. Internal and external computer hardware, such as disk drives and monitors, can be connected to the computer's motherboard by means of these expansion cards. As we learned in earlier lessons, the expansion buses connect to an external data bus. All devices are connected to the same communication bus. In this lesson, we look at how the computer keeps track of each device and controls the flow of data.

I/O Addresses

The bus system establishes a connection between the CPU and expansion devices and provides a path for the flow of data. Everything in a computer, whether hardware or software, requires a unique name and address for the CPU to be able to identify what is going on. Bus-mastering devices might seem to get around this requirement, but they have their own controllers that track local traffic and "talk" to the CPU as needed.

I/O addresses are patterns of 1s and 0s transmitted across the address bus by the CPU. The CPU must identify the device before any data is placed on the bus. The CPU uses two bus wires the IOR (Input/Output Read) wire and the IOW (Input/Output Write) wire to notify the devices that the address bus is not being used to specify an address in memory, but rather to read to or write from a particular device. The address bus has at least 20 wires. However, when the IOW or IOR wire has voltage, only the first 16 wires are monitored.

To allow communication directly between the CPU and a device, each device responds to unique patterns or code built into it. If the CPU needs to check the error status of a hard disk drive controller, for instance, it activates the IOW wire and puts the correct pattern of 1s and 0s onto the address bus. The controller then sends back a message describing its error status.

The following table lists standard PC I/O port address assignments.

PC/XT Port	Used By	PC/XT Port	Used By
000h-00Fh	DMA chip 8237A	2F0h-2F7h	Reserved
020h-021h	PIC 8259A	2F8h-2FFh	COM2
040h-043h	PIT 8253	300h-31Fh	Prototype adapter

060h-063h	PPI 8255	320h-32Fh	Hard disk controller
080h-083h	DMA page register	378h-37Fh	Parallel interface
0A0h-0AFh	NMI mask register	380h-38Fh	SDLC adapter
0C0h-0CFh	Reserved	3A0h-3AFh	Reserved
0E0h-0EFh	Reserved	3B0h-3BFh	Monochrome adapter/ parallel interface
100h-1FFh	Unused	3C0h-3CFh	EGA
200h-20Fh	Game adapter	3D0h-3DFh	CGA
210h-217h	Extension unit	3E0h-3E7h	Reserved
220h-24Fh	Reserved	3F0h-3F7h	Floppy disk controller
278h-27Fh	Parallel printer	3F8h-3FFh	COM1
AT Port	**Used By**	**AT Port**	**Used By**
000h-00Fh	First DMA chip 8237A	278h-27Fh	Second Parallel interface
020h-021h	First PIC 8259A	2B0h-2DFh	EGA
040h-043h	PIT 8253	2F8h-2FFh	COM2
060h-063h	Keyboard controller 8042	300h-31Fh	Prototype adapter
070h-071h	Real-time clock	320h-32Fh	Available
080h-083h	DMA page register	378h-37Fh	First parallel interface
0A0h-0AFh	Second PIC 8259A	380h-38Fh	SDLC adapter
0C0h-	Second DMA chip	3A0h-	Reserved

0CFh	8237A	3AFh	
0E0h-0EFh	Reserved	3B0h-3BFh	Monochrome adapter/parallel interface
0F0h-0FFh	Reserved for coprocessor 80287	3c0h-3CFh	EGA
100h-1FFh	Available	3D0h-3DFh	CGA
200h-20Fh	Game adapter	3E0h-3E7h	Reserved
210h-217h	Reserved	3F0h-3F7h	Floppy disk controller
220h-26Fh	Available	3F8h-3FFh	COM1

I/O addresses have several important characteristics to remember:

- I/O addresses have 16 bits; they are displayed with a hexadecimal number.
- By convention, the lead 0 is dropped (because all I/O addresses have it).
- Hexadecimal I/O addresses must use capital letters; they are case sensitive.

Interrupt Request

The I/O address and the address bus establish a method of communication. The next step is to prevent multiple devices from "talking" at the same time. If the CPU needs to communicate with a device, BIOS routines or device drivers can use I/O addresses to initiate conversations over the external data bus.

Controlling the flow of communication is called *interruption*. Every CPU has a wire called the INT wire. If voltage is applied to the wire, the CPU interrupts what it is doing and attends to the device. For example, when a mouse button is pressed, the CPU attends to the interrupt request, invoking the necessary BIOS routine to query the mouse.

Because the CPU has only one interrupt wire and must handle many peripheral devices, a specific type of chip, called the 8259 chip, is present on the system to help the CPU detect which device is asking for attention. Every device that needs to interrupt the CPU is provided with a wire called an *IRQ* (*interrupt request*). If a device needs to interrupt the CPU, it goes through the following steps:

1. The device applies voltage to the 8259 chip through its IRQ wire.

2. The 8259 chip informs the CPU, by means of the INT wire, that an interrupt is pending.
3. The CPU uses a wire called an INTA (interrupt acknowledge) to signal the 8259 chip to send a pattern of 1s and 0s on the external data bus. This information conveys to the CPU which device is interrupting.
4. The CPU knows which BIOS to run.

Some important facts to remember about IRQs include the following:

- IRQ 2 and IRQ 9 are the same IRQ.
- Three IRQs are hardwired (0—system timer, 1—keyboard controller, and 8—real-time clock).
- Four IRQ assignments are so common that no computer or device manufacturer dares to change them for fear their devices will cause conflicts (6—floppy disk controller, 13—math coprocessor, 14—primary IDE controller, and 15—secondary IDE controller).
- Four IRQs default to specific types of devices but can be changed: IRQ 3—COM2 and COM4, 4—COM1 and COM3, 5—LPT2, and 7—LPT1. (See table that follows.)
- The rest (IRQs 2/9, 10, 11, and 12) are not specific and are available for use.

IRQ	Function	Available for Change
IRQ 0	System timer	No
IRQ 1	Keyboard controller	No
IRQ 2/9	Available	Yes
IRQ 3	COM2, COM4	Usually
IRQ 4	COM1, COM3	Usually
IRQ 5	LPT2	Usually
IRQ 6	Floppy disk controller	No
IRQ 7	LPT1	Usually
IRQ 8	Real-time clock	No
IRQ 10	Available	Yes
IRQ 11	SCSI/available	Yes

IRQ 12	Available	Yes
IRQ 13	Math coprocessor	If there is no math coprocessor
IRQ 14	Primary IDE controller	No
IRQ 15	Secondary IDE controller	Usually

Setting IRQs

Just like I/O addresses, IRQs can be set using hardware, software, or a combination of both. The best way to ensure that no two devices share the same IRQ is to document the IRQs for each device you install in a computer and file that documentation in a location in which you can find it easily if it is needed. As an example, suppose one of your customers has recently installed a sound card that now locks up when a parallel-port tape backup unit is used on the system. This strongly indicates an IRQ conflict. You need merely to check the sound card and the tape backup IRQ settings you have on file and change one if necessary.

Direct Memory Access (DMA)

The CPU runs the BIOS, operating system, and applications, as well as handling interrupts and accessing I/O addresses. This requires the CPU to move a lot of data. This data movement is necessary, requiring considerable CPU power and time, but it is also very simple. Moving data is a waste of the CPU's resources. To reduce this waste, another chip is installed to work with the system CPU. It is called a *direct memory access (DMA)* chip. The only function of the DMA chip (the 8237 chip) is to move data. It handles all the data passing from peripherals to RAM and vice versa.

DMA transfers are not automatic. Hardware and device drivers must be designed to take advantage of this chip. Originally, DMA was used only to transfer data between floppy disk drives and RAM; early computers had only four wires and one DMA chip. Any device requiring DMA had to send a request, just like an IRQ. DMA channels use the same rules as IRQs. The floppy disk drives on all computers use DMA channel 2.

Setting DMA Channels

Fortunately, not many devices use DMA. Sound cards, a few SCSI controllers, and some CD-ROM drives and network cards do require DMA. Just as with IRQs and I/O addresses, DMA can be set by means of either hardware or software. However, manufacturers started using DMA for devices other than the floppy disk drive only recently. As a result, almost all devices set DMA through software (although some still

use jumpers). If two devices share the same DMA channel and "talk" at the same time, the computer will lock up. The following table provides DMA channel assignments.

DMA Channel	Function
0	Available
1	Available
2	Floppy disk controller
3	ECP parallel/available
4	First DMA controller
5	Second sound card
6	SCSI/available
7	Available

COM and Ports

IBM created preset combinations of IRQs and I/O addresses for serial and parallel devices. These preset combinations are called *ports*. The word "port" simply means a portal or two-way access. The preset combinations are called COM ports for serial devices and LPT ports for parallel devices. The purpose of a port is to make installation easier. Modems and printers, therefore, do not require IRQ or I/O settings. When assigned to an active port (as long as no other device is using that port), they will work. The following table lists standard ports.

Port	I/O Address	IRQ
COM1	3F8	4
COM2	2F8	3
COM3	3E8	4
COM4	2E8	3
LPT1	378	7
LPT2	278	5

Installation Problems with COM Ports

Assume you have a modem set to COMl. You buy a network card that comes out of the box with a default setting of IRQ 4. You realize the network card and the modem will conflict, and the computer will lock up. What do you do?

You will have to change the IRQ on one of the devices. The network card is probably the best choice, because the modem is installed and already working.

COM Ports

The original 8088-based IBM PCs were equipped with two serial ports: COMl, set to IRQ4, and COM2, set to IRQ3. While those two IRQs are still the standard for COM ports 1 and 2, many BIOS routines will allow different IRQ assignments or even allow an unused port to be disabled. Because of the limited number of IRQ addresses available, any additional COM ports would have to share IRQs with existing ports. COM3 shared the interrupt of COM1 (IRQ4), and COM4 shared the interrupt of COM2 (IRQ3). To enable use of these additional ports, COM3 was assigned I/O address 3E8-3EF, and COM4 was assigned I/O address 2E8-2EF. This sharing was possible because the IRQ-sharing devices would be unlikely to use them at the same time.

Today we have many other ways of adding printers and other peripherals to PCs, but such conflicts can still be a problem with modems and UPS (uninterruptible power supply) devices that might need simultaneous access.

LPT Ports

LPT ports are for parallel data connections. The name is derived from their original use with printers (LPT—line printer). The original IBM standard LPT port did not provide bidirectional communications (talkback) and was designed solely for one-way data streams to a printer. The standard addresses are IRQ 7 for LPT1 and IRQ 5 assigned to LPT2, if it is present. IRQ 5 quickly became the favorite for devices like sound cards and other add-ons. This situation is easing as USB connections replace many of the parallel designs.

Installing Expansion Cards

The rules for installing expansion cards are simple:

- First read the manual.
- Document addresses and DMA and IRQ settings for any non-Plug and Play device.
- Keep the IRQs, DMAs, and I/O addresses unique.

Windows 95, Windows 98, and Windows 2000 support Plug and Play. In most cases, you can insert a Plug and Play card into the proper type of expansion slot and turn on the computer. Windows will find the card and guide you through the setup. The savvy computer professional documents and keeps track of the IRQ, DMA, and I/O addresses, in case a conflict arises with a Plug and Play device on the system.

Windows 95 and Windows 98 use Hardware Properties, under the System Information/Device Manager option in the Control Panel, which does a good job of identifying (and allowing) changes to these settings. For Plug and Play to work, the computer must have a Plug and Play BIOS, and the operating system and the device card must be Plug and Play-compliant.

Chapter 11

The Display System, Monitor & Adapter

OBJECTIVES

Monitors

Display Adapters

Choosing & Troubleshooting Display Systems

Monitors

Early personal computers employed text-based displays, offering green, white, or amber characters against a black background. Today the average PC monitor can provide life-like colors and reproduce images of near-photographic quality. This dramatic change is the result of radical improvements in monitors and the display-adapter technology that drives them. This chapter discusses how these devices work in unison to provide an acceptable display.

Basic Monitor Operation

A monitor operates fundamentally like a TV set, except that it is designed to receive signals from a card in the PC, rather than a broadcast signal. The quality of a monitor's display is influenced by a variety of design factors and the features of its companion adapter card.

Repairing the inside of monitors is a job more in the realm of a TV repairman than a computer technician. Monitors generally carry warnings that they contain no user-serviceable parts for good reason. Although we discuss the inner workings of monitors in this chapter, that should not be taken as an invitation to probe inside them, where the risk of serious electrical shock is high.

The Cathode-Ray Tube

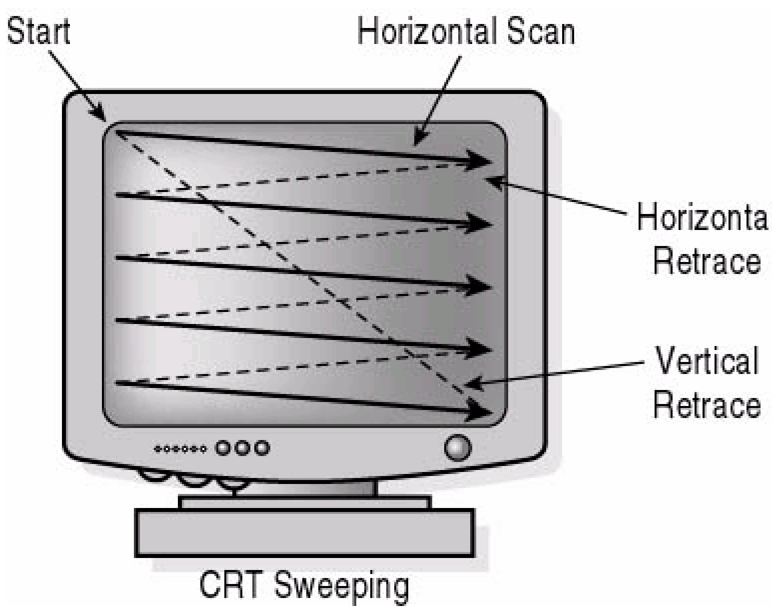
The cathode-ray tube (CRT) is the main component of a traditional monitor. The rear of the CRT holds a cylinder that contains one or more *electron guns*. Most color monitors have three guns in back—one for each of the colors red, green, and blue. This combination (usually referred to as RGB) allows the visual production of all colors.

The wide end of the CRT is the display screen, which has a *phosphor coating* (a substance that can emit light when hit with radiation). When active, the guns beam a stream of charged electrons onto the phosphorus coating. When the coating is hit with the right amount of energy, light is produced in a pattern of very small dots. This same technology is used in X-ray imaging, oscilloscopes, and other CRT devices. Similarly, monitors emit X-radiation. There is one dot for each primary color (RGB), and the dots are grouped in patterns close together. The name for a collection of all dots in a specific location is a *pixel* (which stands for picture element).

Image Formation and Refresh Rates

The image on the screen is not painted all at once. The stream is directed in rows, usually starting in an upper left corner. A series of *rasters* (a rectangular pattern of lines) are drawn down the face of the screen until the beam reaches the lower right, whereupon the process starts over. The persistence rate must hold for long enough to allow formation of a complete image, but not so long that it blurs the dots painted in the next pass.

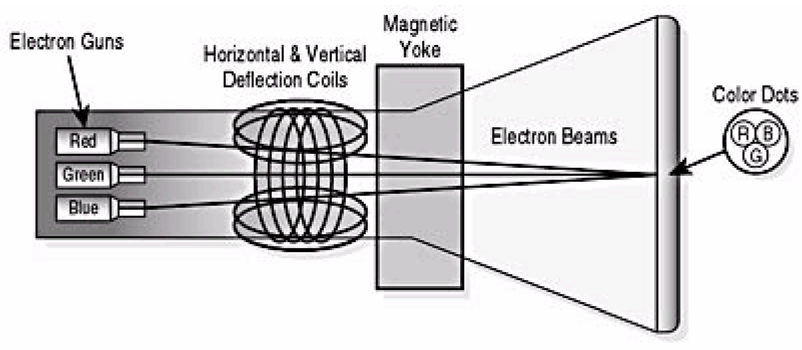
These raster passes take place very quickly. The time required to complete a vertical pass is called the *vertical refresh rate (VRR)*; the time required to pass once from left to right is known as the *horizontal refresh rate (HRR)*. Generally speaking, faster is better. Too slow a vertical rate can cause flicker, which is not only annoying, but can lead to eye strain. The larger the CRT, the faster the refresh rate must be to cover the entire area within the amount of time needed to avoid flicker. At 640 x 480 resolution, the minimum refresh rate is 60 Hz; at 1600 x 1200, the minimum is 85 Hz. The refresh rate, shown in the picture below, is produced by both the monitor and display adapter.



Horizontal and Vertical Refresh Rates

Early monitors had fixed refresh rates. In 1986, NEC introduced the first *multifrequency* monitor that could automatically adjust the refresh rate to take advantage of the highest rate supported by the display adapter of that time. NEC used the term "multiSync" to describe its line of multiple frequency monitors. Today, this feature is standard on most monitors.

The direction and point of contact of the electron stream on the phosphor display are determined by deflection coils coupled with a series of magnetic fields generated by a ring of electromagnets placed around the narrow end of the tube. This collection is called the *yoke,* because it forms a yoke around the tube.



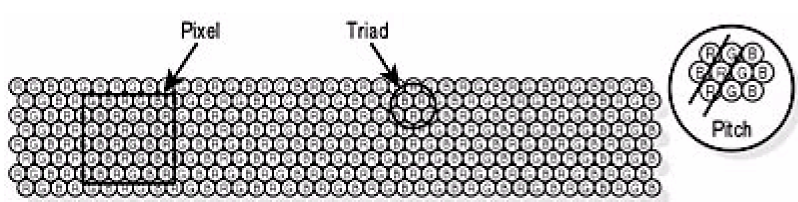
Cathode-ray tube (CRT)

Screen Resolution and Pitch

The term *resolution* refers to the degree of detail offered in the presentation of an image. The method of measurement varies, based on the medium—photographic lenses, films, and paper are measured using lines per inch, while computer-monitor manufacturers express resolution in pixels per inch. The greater the number of pixels per inch, the smaller the detail that can be imaged, and, consequently, the sharper the picture.

Monitor resolution is usually expressed as "a" x "b" where "a" is the number of horizontal pixels, and "b" is the number of vertical pixels. For example, 640 x 480 means that the monitor resolution is 640 pixels horizontally by 480 pixels vertically. Modern monitors usually offer a variety of resolutions, with different refresh rates. Price and quality should be compared at the maximum for both, along with two other factors, dot pitch and color depth (the latter is covered in the next lesson).

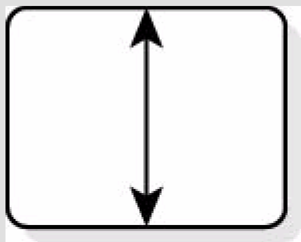
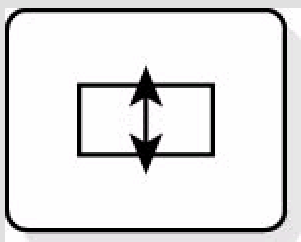
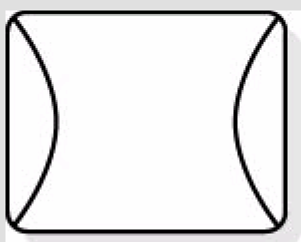
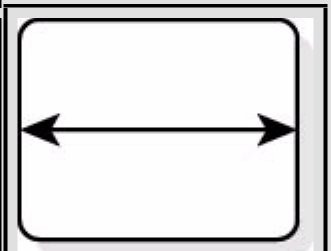
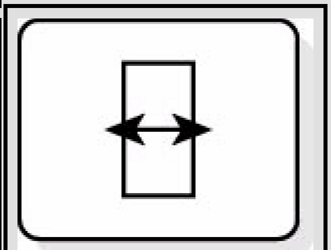
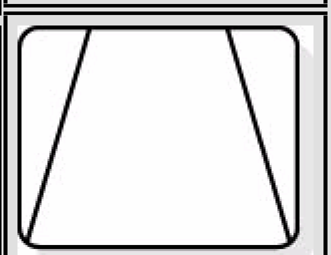
Dot pitch is a term used to define the diagonal distance between the two closest dots of the same color, usually expressed in millimeters.

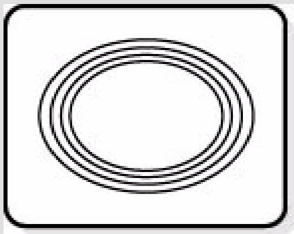


Color dots on a monitor

Tuning the Monitor's Display

In most cases, the monitor must be adjusted for a proper picture when the screen resolution or refresh rate is changed or a new display card is added to the system. The following table lists typical monitor adjustments.

	**Vertical height**	Allows adjustment of the top and bottom of the image. Image height changes.
	**Vertical center**	Allows adjustment of the location of the image between the top and bottom of the active viewing area.
	**Pincushion**	Adjusts the center of the image (vertically) to eliminate or reduce bowing in or out of the display image.
	**Horizontal width**	Allows adjustment of the size of the image on the horizontal axis.
	**Horizontal center**	Allows the image to be horizontally centered.
	**Keystone**	Allows adjustment of the top and bottom edge widths so that the image is square.

	**Degaussing**	Available on newer, larger monitors. Demagnetizes the CRT to prevent an electron beam from bleeding over to an adjacent dot, causing shadowing and/or loss of color control.

Some monitors offer advanced adjustment screws that allow the user, without opening the monitor shell, to tune the settings that are usually available only with interior controls. Adjust these tools only if you understand the process and after you have reviewed the instructions in the product's manual. Any internal repairs or adjustments to the monitor that require opening the shell should be left to properly trained technicians with the appropriate tools and training.

CAUTION

Working inside a monitor is dangerous. The high voltages inside a monitor can cause *sudden death or serious injury*. It's important to realize that monitors contain their high-voltage charge even when they're not operating or even plugged in. A monitor can hold its charge for days. Before you can work safely inside the monitor, such energy must be discharged. Remember that this is often best left to someone who specializes in repairing monitors.

Monitor Maintenance

Monitor care and troubleshooting are usually simple tasks. Here are some general guidelines to follow:

- Make sure the enclosure is properly ventilated. Covering the opening on the case can lead to overheating. Dust the unit at regular intervals.
- Clean the face of the CRT gently: Follow the instructions in the product manual. In most cases, this means dusting the glass with a clean soft cloth. Do not use window cleaners that contain solvents on the unit.
- Make sure that all driver settings are kept within the operating guidelines of the product. Never operate at higher resolutions or refresh rates than those specified by the vendor, and stay within the limits of the display adapter.
- Use any automatic energy-conservation features supported by the hardware and operating system. Employ a screen saver on older models that lack energy-saving features. If possible, do not turn the monitor on and off more than twice a day.
- When a monitor fails to operate or produces an improper image, check the following:
    1. Check all cables, including power and display.

2. Check the front panel controls. Make any appropriate minor adjustments that are needed.
3. Check and, if needed, reinstall the display drivers. Make sure all settings are within the required limits. Reinstall by returning to a plain 16-color, VGA display mode and adding resolution; then increase the refresh rate.
4. Try another display adapter; then, if the problem is still unresolved, try another computer.
5. If the monitor still shows problems, refer to a specialist for further tests.

Display Adapters

The monitor is only half of a computer's display system; it must be matched to a display adapter (also commonly referred to as a graphics adapter, video card, or video controller). This lesson discusses the different types of display adapters and design issues that affect quality and performance.

Evolution of the Display Adapter

The display adapter has gone through several major evolutions as the nature of PC computing has changed from simple word processing and number crunching, to the graphics-intensive world of Windows and multimedia.

The First PC Display Cards

The two "official" video cards for the early 8088-based IBM personal computers (the PC and XT) were matched to the limited capabilities of the early monitors. The *Monochrome Display Adapter (MDA)* offered a simple text-based monochrome display. This adapter produced an 80-character-wide row of text at a resolution of 720 x 350 pixels. Shortly after that, the *Color/Graphics Adapter (CGA)* card appeared. It provided up to four "colors" (actually, just different intensities of the monitor's active color: amber, green, or white). In four-color mode, CGA provided a resolution of 320 x 200 pixels. Using just two colors allowed a resolution of 640 x 200 pixels.

With the release of the *Enhanced Graphics Adapter (EGA)* card, the IMB PC AT became the first PC really able to use color. This adapter was an improved version of CGA, offering a top resolution of 640 x 350 with 16 colors in text-only mode, and 640 x 200 with two colors in graphics mode.

Memory and the Arrival of the Display Coprocessor

A brief digression to explain pixel depth and video memory demands will help you understand what follows. Both the MDA and CGA adapters were equipped with 256 KB of *DRAM (dynamic random access memory)*. The amount of memory on a display card determines the amount of color and resolution that it can image and send to the monitor. As the desire for better graphics and color displays increased, so did the complexity of graphics cards and with them, memory requirements and cost.

The MDA card was a 1-bit device. In other words, each pixel used 1 bit, valued either 0 or 1 to represent whether a given position on the screen (a pixel) was on or off. To represent colors or shades of gray, a card must use memory to describe color and intensity. This attribute of the display, measured in bits, is known as *color depth*. Color

depth multiplied by resolution (the number of horizontal pixels multiplied by the number of vertical rows on the screen) determines the amount of memory needed on a given display adapter. The adapters that followed the EGA cards to market all offered more colors and, very quickly thereafter, higher resolution, use of graphical interfaces and lots of color, with the advent of the graphical user interface (GUI), all that changed.

The new generation of display cards started the practice of including their own display coprocessors on-board. Coprocessors, which have their own memory, are tuned to handle tasks that would usually slow down the PC, and many display cards use bus mastering to reduce the amount of traffic on the system bus and to speed display performance. Video coprocessing is also called "hardware acceleration."

These new graphics chips were designed to do one thing: push pixels to the screen as efficiently as possible. Today, high-performance graphics adapters are the norm. While there is no longer a mad rush to market, the graphics coprocessor is a key element of fast Windows performance. Next, we return to our review of standards and see how the industry progressed to today's world of high-speed, full-color computing.

The Advent of Advanced Display Systems

Graphics artists, engineering designers, and users who work with photorealistic images need more than a coarse, 16-color display. IBM offered a short-lived and very complicated engineering display adapter, the *Professional Graphics Adapter (PGA)*. It required three ISA (Industry Standard Architecture) slots, and provided limited three-dimensional manipulation and 60 frames-per-second animation of a series of images. It was also very expensive and a dismal failure in the marketplace.

The reason was the advent of the *Video Graphics Array (VGA)* standard. All the preceding cards were digital devices; the VGA produced an analog signal. That required new cards, new monitors, and a 15-pin female connector. It allowed developers to produce cards that provided the user with up to 262,144 colors and resolutions up to 640 x 480.

The VGA card quickly became commonplace for a PC display system, and the race was on to produce cards with more colors, more resolution, and additional features. VESA (Video Electronics Standards Association) agreed on a standard list of display modes that extended VGA into the high-resolution world of color and high photographic quality we know today. The standard is known as SVGA (Super VGA). The SVGA sets specifications for resolution, refresh rates, and color depth for compatible adapters. On Pentium and later PCs, an SVGA adapter is the standard for display adapters. The minimum resolution needed for SVGA compatibility is 640 x 480 with 256 colors, and most modern adapters usually go far beyond that.

High-Color, True-Color, Photorealism, and Multimedia Displays

The SVGA specification for 256 shades of gray is in the basic SVGA specifications for true photographic reproduction of monochrome images; it's the number of shades that the human eye expects in a grayscale photo. Color requires the same number of shades for each color in the image to achieve the same level of visual realism. To get 256 shades requires an 8-bit memory address system inside the card, $2^8 = 256$. In the early days of SVGA, vendors worked to increase color quality without significantly increasing cost.

In color mode, an 8-bit card can't display all the colors in a full-color picture, so a LUT is used to figure the closest match to a hue that can't be represented directly. Although this method isn't ideal, it was for several years the state of the art on desktop PCs. Then, early in the age of the 486 processor, came the 16-bit SVGA card, which allowed approximately 64,000 colors. More bits require more memory, more processing requirements, a bigger LUT, and more money. These cards were tuned to be used with bigger monitors, 15 to 17 inches at 800 x 600 or 1024 x 768 resolution. The new systems were too expensive for average users, but graphics professionals and power users generated a large enough market to fuel development.

True Color Arrives

The SVGA adapters were a stepping stone; the growing popularity of Microsoft Windows and scanners pushed the demand for cards that could deliver color of photographic quality. Soon after, stand-alone products that offered both SVGA resolution and true-color operation arrived. These adapters, known as true-color or 24-bit displays, come with coprocessors, lots of memory, and in true color mode have 256 shades (8-bits) available for each of three colors: red, green, and blue. By mixing them, the system can display 2^{24} colors.

Multimedia cards are equipped with a TV tuner, letting the owner view TV programs on the monitor, or watch DVD (digital video disc) movies on a PC. One reason for the dramatic lowering of prices and added features stems from the mass production of the coprocessors, which reduced their cost to the manufacturer; another factor is the decreasing cost of video memory.

Video Memory

As mentioned earlier, the amount of memory on a display adapter is a major factor in determining the screen resolution and color depth that the card can manage. Just as with system RAM, the video memory must be able to operate at a speed that can keep up with the processor, and the demands of the system clock. If the display adapter is too slow at updating the image on the monitor, the user is left waiting or is presented with jerky mouse movements and keystrokes that appear in delayed bursts rather than as typed.

Fast Page DRAM

Early video cards used *fast page-mode (FPM) DRAM*, a series of chips that were basically the same as the RAM used on the early PC's motherboard. This memory form was fine for MDA and CGA cards, and even the 8514/A, but with the higher resolution, increasing pixel depth, and faster refresh rates of VGA displays and beyond, vendors sought improved memory models to get the most performance out of their video coprocessors.

VRAM

Enter dual-ported memory in the form of *VRAM (Video RAM)*. It can read and write to both of its I/O ports at the same time. It allows the processor to talk to the system bus and the monitor simultaneously: fast, but very expensive. VRAM showed up in the best cards, but vendors wanted a low-cost option as well. Some vendors just used FPM (fast page-mode) DRAM, leaving the user to discover that, at high resolution, the display was too slow for efficient operation. These cards did sell well in the low-end market, as they allowed the budget-minded user to operate in low-color modes for most tasks, switching only to higher color depth for projects that required high-color or true-color mode. Users who regularly worked in high-color or true-color mode often would quickly consider an upgrade.

EDO DRAM

An alternate is *EDO (Extended Data Out) DRAM*, which can begin reading a new set of instructions before the preceding set of instructions has been completed. This is a common form of system DRAM that boosts performance to about 15 percent above conventional DRAM.

WRAM

WRAM (Window Random Access Memory, unrelated to the Microsoft operating system) is a high-speed variant of VRAM that costs less to produce and boosts performance by about 20 percent beyond regular VRAM. VRAM and WRAM have become the standard memory types for high-end display adapters.

SGRAM

The mid-range display market makes use of *SGRAM (Synchronous Graphics RAM)*. As the name implies, it is tuned to the graphics-card market, offering faster transfers than DRAM, but not as fast as VRAM and WRAM.

MDRAM

Multibank DRAM (MDRAM) is the final stop on our tour of memory acronyms. It uses interleaving (the dividing of video memory into 32KB parts that can be accessed concurrently) to allow faster memory input/output to the system without expensive dual-porting. It is also a more efficient type of chip that is practical to produce in sizes smaller than a full megabyte. A vendor can save money by just buying the amount needed to actually draw the screen. This saves about 1.75 MB per card for a resolution of 1024 x 768.

The table below lists the standard memory requirements for the most common resolutions and pixel depths used today. As pointed out in the previous paragraphs, keep in mind that the minimum amount of memory for MDRAM is usually less than for other types of RAM. Some graphics cards offer additional memory, and even incorporate different types of RAM on the same card. In such cases, some of the memory might be used for features other than merely imaging the picture to be sent to the CRT in pixels.

Display Drivers

Minimum Memory Requirements for Common Display Resolutions			
Screen Resolution	**8-Bit (256 color)**	**16-Bit (65-KB color)**	**24-Bit (16.7 Million Color)**
640 x 480	512 KB	1 MB	1 MB
800 x 600	512 KB	1 MB	2 MB
1024 x 768	1 MB	2 MB	4 MB
1280 x 1024	2 MB	4 MB	4 MB
1600 x 1200	2 MB	4 MB	6 MB

Text-based adapters under MS-DOS don't need software drivers to interface between the operating system and the image on the screen. Windows, OS/2, and other graphics-rich environments do need drivers. In addition, controls are needed to adjust the refresh rate, resolution, and any special features the card offers. These needs are handled by the use of display drivers, a software layer that marries the card and monitor to the operating environment.

When installing a new card or operating system for a client, be sure to check the manufacturer's Web site for the latest display drivers. Not only will you reduce the

likelihood of problems in using the new addition, but you will find that most new cards incorporate setup routines that can make quick work of getting a new display running.

Choosing and Troubleshooting Display Systems

Matching the components of the display system to the needs of the computer owner is critical to user satisfaction when upgrading or purchasing a system. Knowing the basic steps to troubleshooting a display is key to effecting a quick repair when a component fails to operate properly. This lesson sets out the basic steps to follow in selecting the right class of display hardware for a client, proper care, and troubleshooting common problems.

The CRT size determines the maximum viewable area. For users who will work in only one program at a time, or who don't need high-resolution, a basic monitor should suffice. Graphics intensive applications and multitasking call for larger monitors with faster refresh rates, and display cards to match. Multimedia systems can benefit from cards that offer TV out (usually in the form of an RCA jack that lets the signal be displayed on a regular TV set using the NSTC format), TV tuner, and hardware DVD acceleration.

In recommending a display system, start with your customers' needs, followed by their preferences, and match the two as closely as possible to the available budget. Keep in mind that the display adapter is only part of the equation. Cost can be contained by using a smaller monitor or by accepting slower refresh rates. Cutting the cost excessively can leave clients with a display that does not support the tasks they perform, or that might lead to user eyestrain from the flicker that occurs at slower-than-acceptable refresh rates for the selected resolution.

Troubleshooting Display Systems

When MDA cards were standard, display systems gave technicians few problems. If the cable was properly attached and the monitor was working, the user got a picture. Today, the wide range of card options and the mix of resolutions, refresh rates, and operating systems lead the user to require help with displays more often.

In spite of the increasing complexity of display systems, most problems can be traced to a few common sources: cables improperly connected or damaged; lack of power; improper monitor adjustment; corrupt or incorrect drivers; and memory conflicts with other components. The following checklist can help you troubleshoot the common display problems you are likely to encounter:

- Cables: Verify that both the power and monitor-display adapter cables are properly attached. Failure to attach them properly can lead to having no picture at all or to an erratic image with incorrect colors. If the monitor cable has been removed and reseated, bent pins could be the problem. Make sure power is reaching both the PC and the monitor, and that they are turned on.

- Make sure that the adapter is properly seated in the expansion slot.
- Boot the system. If you get an image during the POST (power-on self test), but the computer does not load the operating system, suspect memory or driver problems. The same is true if the system repeatedly hangs during Windows operation. Try working in Safe mode. If that succeeds, reinstall the drivers and use Device Manager in the Control Panel System applet to resolve any hardware or memory conflicts.
- Reset the card to the 640 x 480 in 16-color VGA mode at the 60-Hz refresh rate. If the card works in Normal Mode, in Windows, at these settings, yet fails at higher resolution, color depth, or refresh rates—check the drivers and the capabilities of the display components.

Chapter 12

Printers, Modem and Cables

OBJECTIVES

Printers

Modems

Cables & Connectors

Printers

Printers are considered standard PC components; they are often bundled with computers and sold to consumers as part of a complete package. The most common add-ons, printers are manufactured in several popular forms. Like other devices, each type has unique advantages and disadvantages. This lesson covers all types and aspects of printers.

Printer Basics

Printers have virtually replaced the typewriter in the contemporary office. The simplest printers are designed for the bargain-seeking home user, while the most complex are designed for production use, producing over 60 pages per minute and offering features like collating, stapling, and internal two-sided printing.

When evaluating printers, you should keep the following issues in mind:

- **Printer resolution:** Resolution is usually measured in dots per inch (dpi). This indicates the number of vertical and horizontal dots that can be printed; the higher the resolution, the better the print quality.
- **Speed:** This is usually given in pages printed per minute, where the page consists of plain text with five percent of the printable page covered in ink or toner.
- **Graphics and printer-language support:** If the device is used to print graphics, it should support one or more of the popular printer languages, such as Adobe PostScript and Hewlett Packard's LaserJet PCL (Printer Control Language).
- **Paper capacity:** The number and types of paper trays available, the number of pages that can be placed in them, and the sizes of pages that can be printed all vary widely among printers. Some smaller units hold as few as 10 sheets, while high-volume network printers hold several reams in different sizes. Some printers can also be set to automatically choose which tray to use based on the type of paper best suited for a job.
- **Duty Cycle:** This is the number of sheets of paper the printer is rated to print per month. It is based on a plain-text page with five percent coverage and does not include graphics.
- **Printer memory:** Laser printers that will be used to print complex graphics and full-color images require larger amounts of memory than those which print simple text only. In many cases, this memory can be added as an option.
- **Cost of consumables:** When comparing the cost of various printers, be sure to calculate and compare the total cost per page for printing, rather than just the cost of a replacement ink or toner cartridge.

Common Printer Terms

You should familiarize yourself with these basic terms used with printer communication:

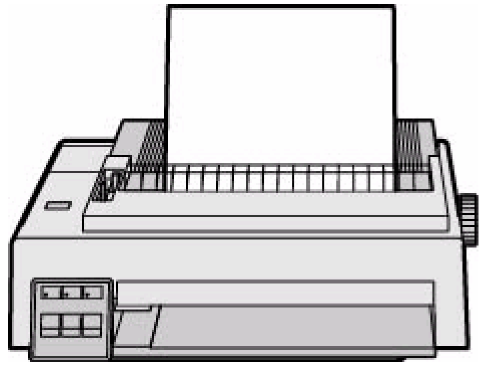
- **ASCII (American Standard Code for Information Interchange):** A standard code representing characters as numbers, used by most computers and printers.
- **Font:** A collection of characters and numbers in a given size, usually expressed in style name and size, and expressed in points (pts.)—for example, Times Roman 10 pt. bold (One point equals 1/72 inch.) Although many people think that bold and italic are variations of the same font, technically they are different fonts. Some printers are sold with limited fonts, such as bold only or no-bold varieties of the typefaces.
- **LPT (Line Printer Terminal) Port:** Term that describes parallel printer ports on a computer.
- **PCL:** Hewlett-Packard's printer-control language for printers.
- **PostScript:** The most common page-description language (PDL). A method of describing the contents of a page as scalable elements, rather than bitmapped pixels on the page. The printer is sent a plain ASCII file containing the PostScript program; the PostScript interpreter in the printer makes the conversion from scale to bitmap at print time.
- **Resolution Enhancement:** Technology that improves the appearance of images and other forms of graphics by using such techniques as modifying tonal ranges, improving halftone placement, and smoothing the jagged edges of curves.
- **Portrait:** The vertical orientation of printing on a piece of paper so that the text or image is printed across the 8.5-inch width of the paper.
- **Landscape:** The horizontal orientation of printing on a piece of paper so that the text or image is printed across the 11-inch width of the paper.
- **Duplexing:** The ability to print on both sides of a page. This cuts operating costs and allows users to create two-sided documents quickly.

Printer Ports

- The standard parallel port uses cable with a 25-pin female connector on one end and a Centronics-compatible D-Shell fitting on the other. You simply connect the Centronics-compatible end of the printer cable to the printer and attach the 25-pin plug to an LPT port on the computer.

Impact Printers

In the early days of PC printing, the most commons forms of printers were dot-matrix and daisy-wheel designs. Both these designs create an impression by striking an inked ribbon with enough force to place ink on the page. In this, they function very much like typewriters. Except for a few special cases, impact printers have been replaced by ink and laser technology.

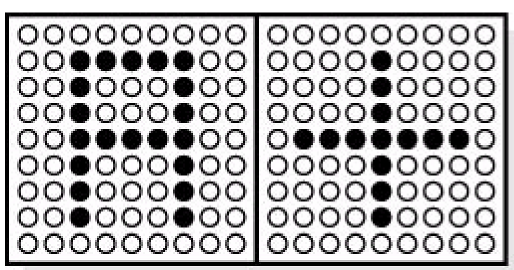


Impact printer

Dot-Matrix Printers

The most popular impact printer is the dot-matrix variety. Dot-matrix printers form characters as raster images on paper by pressing pins onto an inked ribbon, which then is pressed onto paper. Dot-matrix printers use an array of pins (commonly 9 or 24 pins) which are made of stiff wire. The higher the number of pins, the more dots per square inch and the higher the print quality. The pins are held in a print head that travels on a rail in front of a roller that transports the paper. The pins are controlled by electromagnets; dots are created when power is applied to selected electromagnets in the print head, forcing the desired pin away from a magnet in the print head. The pins strike an inked ribbon, which then strikes the paper. As the individual dots are struck, a character is formed.

The dot-matrix printer is still used in some commercial environments where its ability to inexpensively create multipage forms is appreciated. T



Dot-matrix letters formed from dots.

Maintaining a dot-matrix printer is very simple:

- Change the ribbon.
- Keep the printer clean.
- Keep the print head clean.
- Replace the print head if it fails.

Troubleshooting a dot-matrix printer usually requires a reference manual. There are so many printers on the market that no single guide will suffice to help a computer technician troubleshoot all printer problems.

Symptom	**Possible Cause**
Printer does not function at all.	No AC power is getting to printer. Fuse is blown.
Device does not print although power is on.	Printer is not online. Printer is out of paper. Printer cable is disconnected.
Printer won't go online.	Printer is out of paper. (Check connections.)
Paper slips around platen	Paper is not being gripped properly. (Adjust paper-feed selector for size and type of paper.)
Head moves, but does not print.	Ribbon is not installed properly or is out of ink.
Head tears paper as it moves.	Pins are not operating properly. (Check pins; if any are frozen, the head needs to be replaced.)
Paper bunches up around platen.	There is no reverse tension on paper.
Paper has "dimples."	Paper is misaligned, or the tractor feed wheels are not locked in place.
"Paper/Error" indicator flashes continuously.	There is an overload condition.
Printout is double-spaced or there is no spacing between lines.	Printer configuration switch is improperly set. (Make sure it isn't set to output a carriage return or linefeed after each line.)
Printer cannot print characters above ASCII code 127.	Printer configuration switch is improperly set.
Print mode cannot be changed.	Printer configuration switch is improperly set.

Daisy-Wheel Printers

Once popular, the daisy-wheel printer is now a rare form of impact printer. A daisy-wheel device prints characters one at a time using a circular printer element that produces letter-quality output. Each printer element contains characters of only one font, which means that every time you want to change fonts (for example, from Courier to Times Roman), you must manually change the element. In this respect, these printers are similar to some electric typewriters; these devices produced clean type, along with lots of noise and often a great deal of vibration. If a client has a daisy-wheel printer that needs repair or replacement, consider recommending a laser printer instead.

Thermal Printers

A thermal printer heats spots on paper with an array of tiny, fast-acting heating elements. At one time, these printers were an inexpensive alternative to dot-matrix printers for the low-end, low-volume market, even though their paper costs were much higher. Early fax machines, today's non-plain-paper fax machines, and some label makers are examples of thermal printers. For uses other than faxing, this type of printer is rare, too. You might find it used to print simple sales receipts.

Thermal Wax-Transfer Printers

A thermal wax-transfer printer uses heat to bond a wax-based pigment to paper. These printers usually have resolutions of 150 or 300 dpi. At one time, these devices were used for color graphics and photo-reproduction; more recently, they have been replaced by ink-jet and color laser printers, which are less expensive to own and operate.

The bewildering assortment of transport rollers, heating arrays, and electronics make field repair of these printers impractical, except by a specially trained technician equipped with special tools.

Ink-Jet Printers

Ink-jet printers have replaced dot-matrix printers, in the low-end market, and thermal wax printers for the low-end color market. Many computer manufacturers and large computer stores offer an ink-jet printer as part of a computer-system package deal.

Ink-jet printers spray ink onto paper in order to form images. They produce good-quality printing and—compared to dot-matrix and wax printers—they are relatively fast. They also require little maintenance beyond cleaning and ink-cartridge replacement. What makes them so attractive is their ability to easily produce color as well as standard black-and-white images.

IMPORTANT

When installing a new ink-jet cartridge or replacing the cartridge on an existing ink-jet printer, follow the instructions carefully! The cartridge is not just a simple ink container. It must be properly pressurized, and there are sensors on the unit (small metal plates) that must line up with contacts on the cartridge transport. Read the product manual for details.

When recommending an ink-jet printer, the cost of printing as well as the cost of the printer itself should be considered. The cartridges are usually more expensive per page than those for a laser printer.

If a printer fails to operate, the first step in determining the source of the problem is to decide if the problem lies with the printer or with the computer. The best place to start is at the printer, with a visual inspection. Look for simple issues, like a tray out of paper or a paper jam. Most printers have either an LED (light-emitting diode) panel or lights that warn of common problems.

If visual inspection of the printer does not turn up an obvious fault, proceed to the printer's self-test program. In most cases, you can initiate this routine by holding down a specified combination of control keys on the printer (check the owner's manual for diagnostic procedures) while you turn it on. If a test page prints successfully, the problem is most likely associated with the computer, the cabling, or the network. The following table lists some typical problems encountered with ink-jet printers and their possible causes.

Symptom	**Possible Cause**
Power is on but device does not print.	Printer is not online. Printer is out of paper.
Printer won't go online after user has replaced ink cartridge.	Cartridge is installed incorrectly. Printer cable is disconnected.
Printer is plugged in, but all indicator lights are off and the printer appears to be dead.	Check the drive mechanisms and motors for signs of binding. They might need to be replaced. Fuse is blown. (Check the power supply's fuse and replace with one of the same type and rating, if necessary.)
Print head does not print.	Ink reservoirs are empty. (Check the ink supply and replace the ink cartridge as necessary.)
Paper does not advance.	Paper-handling hardware is jammed. (Check the control panel to confirm that the printer is online. If so, you will need to inspect the paper-handling motor and gear train.

	You can do this by setting the printer offline and pressing down the form-feed button.)

Laser Printers

Run the **laser** video located in the **demos** folder on the CD accompanying this book to view a presentation of laser printers.

The laser printer has become the dominant form of computer output device, with models ranging from personal, low-volume desktop printers to behemoths that fill half a room and serve hundreds of users, churning out reams of pages every day.

All laser printers follow one basic engine design, similar to the ones used in most office copiers. They are non-impact devices that precisely place a fine plastic powder (the toner) on paper. Although they cost more to purchase than most ink-jet printers, they are much cheaper to operate per page, and the "ink" is permanent. (Most ink-jet images are, at best, water-resistant.)

Primary Components of a Laser Printer

A laser printer is a combination of mechanical and electronic components. Although the internal workings of the printer generally are not a concern of the average PC technician, you should be familiar with the parts and processes involved in their operation.

Paper Transport

The paper path for laser printers ranges from a simple, straight path to the complicated turns of devices with options such as duplexers, mailboxes, and finishing tools like collators and staplers. The goal is the same for all these devices—to move the paper from a supply bin to the engine where the image is laid on the paper and fixed to it, and then to a hopper for delivery to the user. Most printers handle a set range of paper stocks and sizes in the normal paper path, and a more extensive range (usually heavier paper or labels) that can be sent though a second manual feed, one sheet at a time. When users fail to follow the guidelines for the allowed stocks, paper jams often result.

Logic Circuits

Laser printers usually have a motherboard much like that of a PC, complete with CPU, memory, BIOS, and ROM modules containing printer languages and fonts. Advanced models often employ a hard disk drive and its controller, a network adapter, a SCSI host adapter, and secondary cards for finishing options. When upgrading a printer, check for

any updates to the BIOS, additional memory requirements for new options, and firmware revisions.

User Interface

The basic laser printer often offers little more than a "power on" LED and a second light to indicate an error condition. Advanced models have LED panels with menus, control buttons, and an array of status LEDs.

Toner and Toner Cartridges

To reduce maintenance costs, laser printers use disposable cartridges and other parts that need periodic replacement. The primary consumable is toner, a very fine plastic powder bonded to iron particles. The printer cartridge also holds the toner cylinder, and often the photosensitive drum. The cartridge requires replacement when the level of toner is too low to produce a uniform, dark print. Some "starter" cartridges shipped with a new printer print only 750 sheets or so, while high-capacity units can generate 12,000 or more pages. Keep in mind that what constitutes a "page" is based on a five-percent coverage area, less than a standard letter and far below that of a page of graphics or images.

Photosensitive Drum

The photosensitive drum is a key component and usually is a part of the toner cartridge. The drum is an aluminum cylinder that is coated with a photosensitive compound and electrically charged. It captures the image to be printed on the page and also attracts the toner which is to be placed on the page.

IMPORTANT

The drum should not be exposed to any more light than is absolutely necessary. Such exposure will shorten its useful life. The surface must also be kept free of fingerprints, dust, and scratches. If these are present, they will cause imperfections on any prints made with the drum. The best way to ensure a clean drum is to install it quickly and carefully and leave it in place until it must be replaced.

The Laser

The laser beam paints the image of the printed page on the drum. Before the laser is fired, the entire surface of the photosensitive drum, as well as the paper, are given an electrical charge carried by a pair of fine wires.

Primary Corona

The primary corona charges the photosensitive particles on the surface of the drum.

Transfer Corona

The transfer corona charges the surface of the paper just before it reaches the toner area.

Fuser Rollers

The toner must now be permanently attached to the paper to make the image permanent. The fuser rollers—a heated roller and an opposing pressure roller—fuse toner onto the page. The heated roller employs a nonstick coating to keep the toner from sticking to it. The occasional cycling heard in many laser printers is generated when the fuser rollers are advanced a quarter turn or so to avoid becoming overheated.

Erase Lamp

This bathes the drum in light to neutralize the electrical charge on the drum, allowing any remaining particles to be removed before the next print is made.

Power Supply

Laser printers use a lot of power and so should not be connected to a UPS (uninterruptible power supply) device. The high voltage requirements of the imaging engine and heater will often trip a UPS. In addition to the motors and laser print heads, the printer also has a low DC voltage converter as part of the power package for powering its motherboard, display panel, and other more traditional electronic components.

Drivers and Software

Most laser printers ship with a variety of software that includes the basic drivers that communicate with the operating system, diagnostic programs, and advanced programs that allow full control of all options as well as real-time status reporting. A recent trend in network laser printing is allowing print-management tools and printing to work over the Internet. A user can send a print job to an Internet site or manage a remote print job using a Web browser.

The Mechanics of Laser Printing

Now that you know the major parts of the printer, here's a quick survey of how the laser printer works and the components needed to handle each task. (Some of these tasks occur simultaneously in actual printing.)

Communicating

The following sequence occurs as computer-to-printer communication is established:

1. The operating system sends a request to the printer and is informed that the printer is online and ready to accept data.
2. The PC starts sending data.
3. During the printing process, the printer—if it is able to handle bidirectional communications—informs the computer of any problems encountered while handling the print job so the user can address the complaint. These messages might include an out-of-paper condition, paper jam, or low toner.
4. After the entire job has been sent, the printer acknowledges the receipt of all data and waits for the next request.

Many printers can store more than one job, and network printers often have hard disk drives that can hold common jobs to allow the printer to print without being connected to a PC.

Warming Up

The printer might delay accepting the job or printing the first page while it warms up its rollers and the imaging drum.

Raster-Image Processing

The image (text and graphics) to be made is converted into a series of raster lines that can be drawn much the same way as the image is formed on the PC's monitor. The data is stored in memory, waiting for the send command.

Paper Feeding

The printer moves a sheet of paper from the proper tray onto a series of rollers, through the imaging and fixing areas, and to the output hopper.

Drum Cleaning and Charging

Any residual toner from past jobs is scraped from the printer's photosensitive drum. A fine wire (the primary corona) produces a negative electrical charge across the entire face of the drum. The image is set in raster lines as a series of fine dots on the drum.

Imaging the Drum

The information from the raster-image processor is read from memory and sent to the print engine, one line at a time. The laser sets a positive charge in the areas of the image to be filled with toner.

Transferring Toner to the Drum

A film of fine plastic power is placed on the toner transfer roller, which is turning close to the photosensitive drum. This toner is then attracted to the positively charged areas of the drum.

Transferring Toner to the Paper

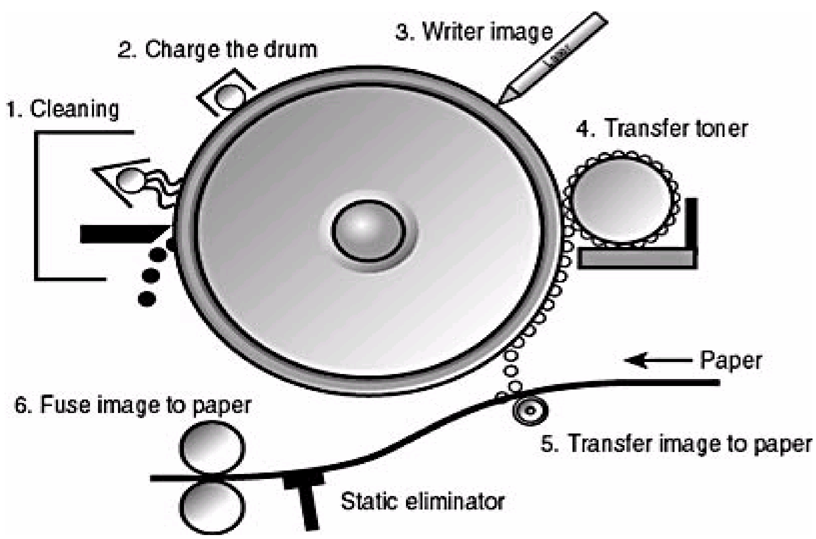
The corona wire places a positive electrical charge on the paper as it moves close to the drum. The toner is attracted to the page, forming an image.

Fusing the Toner

The page passes through a pair of rollers. The roller on the side toward the toner that has been placed on the page is heated just enough to melt the plastic toner particles onto the page without smearing. The roller on the other side supplies the needed pressure.

Finishing and Output

After the toner has been fused to the paper, the next step is usually to transport the page to the output tray. However, if other options—such as a duplexer or collator—are available, the page might be routed through a separate path, based on the options for the current print job, and then sent out to the tray. The picture below shows the process of laser printing.



The laser printing process

Laser Printer Resolution

The quality of a laser printer is directly related to its resolution, given in dpi. Horizontal resolution is determined by how fine a line can be focused on the drum by the laser (the number of dpi across the page); vertical resolution is based on the increment by which the photosensitive drum is turned for each pass of the raster line.

In most cases, resolution is given as a single number, indicating that both the horizontal and vertical increments are the same. The first laser printers offered 300 dpi resolution; printers today commonly offer 600 and 1200 dpi. The higher the number, the sharper the detail and the more memory required to image the page. In general, the human eye cannot distinguish between 600 dpi and 1200 dpi text on bond paper, but the higher resolution does benefit images and drawings by providing a smoother transition between tones and curved lines.

TIP

Many laser printers offer a "toner saver" that uses a lower-resolution draft mode, thereby extending the life of a toner cartridge by placing less toner on each page.

Troubleshooting Laser Printer Problems

Properly installed laser printers are quite reliable when operated and maintained within the guidelines set by the manufacturer. Still, given the combination of mechanical parts, the variety of steps in printing, and the innovative ways some users use the printer, problems do occur. The following table lists a few problems that can be encountered with laser printing and their possible causes.

Symptom	Possible Cause
Ghost images appear at regular intervals on the printed page.	Photosensitive drum is not fully discharged. Previous images used too much toner, and the supply of charged toner is either insufficient or not adequately charged to transfer to the drum.
Light ghosting appears on pages.	Previous page(s) used too much toner; therefore, the drum could not be properly charged for the image (called developer starvation).
Dark ghosting appears on pages.	Drum is damaged.

Page is completely black.	Primary corona, laser scanning module, or main central board has failed.
Random black spots or streaks appear on page.	Drum was improperly cleaned; residual particles remain on drum.
Marks appear on every page.	Drum is damaged and must be replaced.
Printing is too light (appears in a column-like streak).	Toner is low.
Memory overflow error.	Not enough RAM—printing resolution too high.
Characters are incomplete.	Print density is incorrect. (Adjust the darkness setting on the toner cartridge.)
Mass of melted plastic is spit out.	Wrong transparency material is used (see section on transparency, later in this lesson).
Pages are creased.	Paper type is incorrect.
Characters are warped, overprinted, or poorly formed.	There is a problem with the paper or other media or with the hardware. (For media: avoid paper that is too rough or too smooth. Paper that is too rough interferes with fusing of characters and their definition. If the paper is too smooth, it can feed improperly, causing distorted or overwritten characters. For hardware: run the self-test to check for connectivity and configuration problems.)
After clearing a paper jam from the tray, printer still indicates a paper jam.	Printer has not reset. (Open and close the cover.)
Paper continues to jam.	Problem with the pickup area, turning area, and registration (alignment) area. (Look for worn parts or debris.)

Ghosting

The term ghosting is used to describe unwanted images that are produced on the printed page at regular intervals. This usually occurs when the drum is not fully discharging or is being saturated with excess toner. One remedy is to print one or two totally black pages; this will pull the toner off on the drum and onto the paper. If the problem persists, try using a new toner cartridge (if the toner is part of the cartridge assembly). If those steps fail, the printer will require servicing by a trained technician, who is able to adjust the internal settings that regulate toner levels during printing.

Printing on Transparencies

When printing on transparencies, use only materials approved for laser printers. Laser printers generate far more heat than other types of printers—using the wrong material can cause serious damage.

CAUTION

Printing on transparencies can be hazardous to the printer. Improper material might melt and damage the internal components of a laser printer.

Hardware Problems

Most laser printers offer the ability to print a page or more of diagnostic and configuration information. If you suspect a hardware problem, print these sheets. Check for status lights, menu warnings, or error messages. The manual should list steps to be taken in troubleshooting the common problems that are indicated by the printer's display. The variety of error codes that exist, the result of different options on printers, even from the same vendor, makes a detailed listing beyond the scope of this volume and beyond the skills required for the exam. Refer to the printer's manual for details concerning codes for a given printer.

Modems

Modem Basics

A modem is a peripheral device that enables computers to communicate with each other over conventional telephone lines, ISDN cable lines, or even without wires. The word *modem* comes from combining the words MOdulator and DEModulator. In radio parlance, to *modulate* a signal is to change the frequency (FM) or amplitude (AM) of a carrier (fixed signal) by superimposing a code (voice or other information) on top of the signal. The reverse is to *demodulate*, to remove the fixed signal and have the superimposed code remain. For many years, this has been an effective way to communicate over long distances, using both wires and radio waves.

The following table defines some basic terms used with modem communication.

Term	Definition
Baud rate	The number of events, or signal changes, that occur in one second. It was used as an early measurement of how fast a modem can send data, because at that time, modems transmitted data at a speed equal to the baud rate (one bit per cycle). Today's high-speed modems use complex signals to send more data; therefore, data transfer can exceed baud rate. The baud rate is limited by the capability of copper wires to transmit signals.
bps	Stands for bits per second—the speed at which a modem transmits data. Typical rates are 14,400; 28,800; 33,600; and 56,600 bps. These numbers represent the actual number of data bits that can be transmitted per second.
Browser	Software used to explore sites on the World Wide Web using the HyperText Transfer Protocol (HTTP).
Bulletin board service (BBS)	Interactive software that offers the ability log directly onto a remote computer to post and retrieve messages, and to upload and download files. Some BBSs offer the opportunity to chat live with other members who are online at the same time. The Internet has diminished the importance of BBS software.
Download	The act of transferring a file from a remote computer (host)

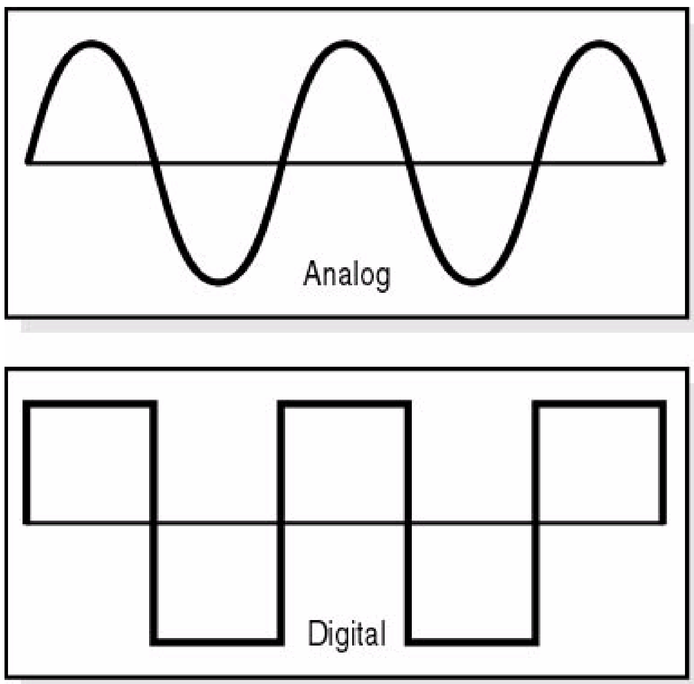
	to a local computer (client). When downloading, you are receiving a file.
DTMF	Stands for dual-tone multifrequency, the technology behind the tones of a touchtone phone.
Internet	A worldwide, online network that links computers by means of the TCP/IP protocol. (See entry for TCP/IP later in this table.)
Internet server	A computer/software combination that provides a gateway and supporting services for linking other computers to the Internet.
IP	Stands for Internet Protocol—the network protocols used to define how data is transmitted on the Internet.
IP address	Stands for Internet Protocol address—a unique, 32-bit address that identifies every network and host on the Internet. (A host is the TCP/IP network interface within the computer, not the computer itself; a computer with more than one network interface card—NIC—can have more than one IP address, one for each card.)
ISDN	Stands for Integrated Services Digital Network, a digital telephone connection like a modem that uses digital links and offers speeds about five times that of an analog modem. The ISDN system is a packet system that can also handle voice communication.
ISP	Stands for Internet service provider—a host computer that you can dial into over a modem to connect to the Internet.
Logging on	The process of sending the appropriate signals and gaining access to a remote computer over a modem or other remote connection.
Modem	A device for converting a computer's digital data stream to and from an analog form so that it can be sent over a telephone line.
Offline	Status of a computer that is not connected to another device over a modem or other telecommunications device.

Offline reader	A program designed to display e-mail messages or other information that has been downloaded to one computer from another.
Online	An active connection between two computers, making possible the exchange of data.
POTS	Stands for Plain Old Telephone Service, a common term that denotes the basic analog phone network as compared to newer digital networks used for packet transfer of data.
Protocol	A set of rules that govern the transfer and verification of data between two or more systems.
Proxy server	A computer/modem/software combination that manages Internet traffic to and from a local area network (LAN) and can provide other features, such as document caching and access control.
Settings (modem)	Configuration required by the telecommunications software in order for data to be transmitted. Usually 8 bits, no parity, 1 stop bit (8N1) will work. Some BBSs require special terminal emulation, a software configuration that mimics the operation of proprietary mainframe terminals like the VT100.
Sysop	The system operator of a BBS (pronounced SIS-op).
TCP/IP	Stands for Transmission Control Protocol/Internet Protocol—the name given to a collection of protocols that were designed in the 1970s for use on the large-scale, mixed-platform environment that became the Internet.
Telecommunications	The ability to transmit data over telephone lines to a remote computer. Often abbreviated as "telecom."
Telecommunications software	An application that allows two computers to communicate with each other. Both computers must use compatible software for communication to take place.
Upload	The ability to transfer a file from a local computer to a remote computer. When uploading, you are sending a file to another computer.

Webmaster	A person who manages a Web site on the Internet's World Wide Web. The Webmaster's duties are analogous to those of a BBS.

Communication

Two basic problems arise from using modems to transmit data. The first problem is that, as we saw in Chapter 4, computers transfer data using 8, 16, or 32 parallel wires or buses, while telephone systems use only two wires. The second problem is that telephone and radio systems use analog signals (based on waveforms), and computers use digital signals, either on or off, as shown in the picture below.



Analog and digital signals

A modem resolves both of these problems by acting as an analog-to-digital (A/D) converter as well as a modulator/demodulator.

Digital Communication

The movement of data from one computer to another over telephone lines is a multistep process. The first step is to convert the data from parallel to serial form. Then the digital information must be broken into uniquely marked packets (this allows the receiving computer to distinguish one byte from another).

Asynchronous Communication

Asynchronous communication is any data transmission that does not link the two devices with a common data clock. This is useful because the length of time between sending a packet and its receipt on the other end can vary between the communicating devices. A signal called a *start bit* is sent at the beginning of each segment, and a signal called a *stop bit* is sent at the end. These let the receiving device note the boundaries (beginning and end) of a transmission packet. Early PC modems were almost always asynchronous devices operating at speeds of no more than 18,000 bits per second.

Synchronous Communication

Synchronous communication sends data blocks at strictly timed intervals that are monitored at both ends. Modems operating at speeds up to 56 Kbps over standard telephone audio lines are usually synchronous devices.

Many protocols are used for PC-to-PC modem communication. Kermit, Xmodem, Ymodem, and Zmodem are four common ones. The following is a summary of how these protocols work:

1. Before a modem sends any data, a communication link must be established. To do this, the modem sends a series of standardized bytes—called sync bytes—to the device it is to communicate with.
2. The modem on the other end receives the sync bytes.
3. The receiving modem perceives that it is receiving sync-byte data and synchronizes with the incoming data.
4. After sending the sync bytes, the sending modem adds a start-of-text (STX) character.
5. The data bytes are sent. The data in synchronous transmission is processed in packets or in blocks of fixed length, depending on the protocol used.
6. Each packet ends with an end-of-text (ETX) character and two error-checking characters called CRC (cyclical redundancy check) characters or BCCs (block check characters).
7. The receiver then responds with an ACK—acknowledgment character—if the data is good, or a NAK—negative acknowledgment—if transmission errors have occurred.

NOTE

In asynchronous communication, the receiving modem does not respond—it just reads the data and acts on it—unless a timing error is reported. In synchronous modes the receiving modem *must* respond.

Parity

Asynchronous communication packets have an optional *parity* bit that is used for error detection. The parity bit is used by the receiving port to verify whether the data is intact or has been corrupted. There are two types of parity:

- **Even parity**: The sending computer counts the 1s in the data part of the packet; if the number of 1s is even, the parity bit is 0—this makes the total number of bits even. If the number of 1s in the data part of the packet is odd, the parity bit is set to 1—again making the total number of bits even. The receiving port counts the data bits and compares its answer to the parity bit. If the two fail to match, an error is reported, and a request to retransmit the packet is passed to the sending computer.
- **Odd parity**: This works in exactly the same way as even parity, except that the total number of bits must be odd.

The use of parity bits is optional. The quality of data transmission and telephone lines has improved to the extent that parity bits are no longer required. However, if data accuracy is critical and/or telephone-line quality is questionable, use parity.

Hardware

Now that we've seen how modems send and receive data, we examine the hardware involved.

Internal Modems

The entire modem and even its serial port can be accommodated on a single expansion card. This configuration offers lower cost than that of an external modem, but is more prone to compatibility problems with either the on-board UART or the COM port IRQs.

USB Modems

Most new PCs offer two universal serial bus (USB) ports, either of which can be used to attach a modem. USB is a hot swap (the device can be added or removed without powering down the PC), Plug and Play interface well suited to this task. To install a modem this way, usually all that is required is to attach a USB cable between the modem and PC, connect the phone-line cable between the modem and a wall jack, and load the modem-driver software from the manufacturer's configuration disk when prompted.

External Analog Modems

The original modems used a pair of cups to cradle a telephone handset over a built-in speaker and microphone; in this way, the modem would send and receive tones acoustically, and the telephone handset would relay the tones. Today, the external modem is usually a rectangular box with a row of status lights on the front, a speaker to give audible feedback, and a number of ports on the back. Two of those ports are telephone jacks—one to connect to the wall line and the other to pass the telephone signal to a phone for regular voice conversations when the modem is not in data mode. A third port on the back of the modem is a serial port using a standard 25-pin RS-232 connector that passes data to and from a serial port on the PC.

ISDN Terminal Adapters

ISDN is an all-digital phone connection that uses special high-quality phone lines to ensure clean, high-speed, data transfers—directly to the user's home or business. Both voice and data are carried by bearer channels (B channels) with a maximum speed per channel of 64 Kbps. A companion data channel (D channel) handles signaling at 16 Kbps (or 64 Kbps, depending on service provided by the carrier).

ISDN connections do not make use of a modem. Instead, a device called a terminal adapter (TA) serves as the interface for both computers and analog phones served in a location. Most small-business and residential customers make use of a TA that has a 25-pin serial connection to attach to a computer serial port and that also provides analog telephone connections for two lines.

ISDN is complicated to install and should be set up using the help of a vendor or the local telephone company. After installation, ISDN functions like a high-speed modem, offering not only faster data transfers, but faster connections to remote ISDN providers such as ISPs. Because each TA unit is completely digital, there is no testing of the nature of the remote source by the hardware to establish the maximum connection rate (as with a modem), and links are typically established in under three seconds.

The RS-232 port (External Modems)

External modems and TAs communicate with their host computers by means of an RS-232 communications port. The RS-232 standard was developed by the Electronics Industry Association (EIA) for low-speed data communication; the standard defines a series of signals that are sent between two telecommunications devices to indicate line and transmission status. The following table shows the most common signals.

RS-232 Cables

RS-232 connections can make use of either 25-pin or 9-pin connectors. On many PCs, the end attaching to a modem or TA has a 25-pin connector, while the PC has a 9-pin connector. The following table presents the layout and signals for both.

Telephone-Line Basics For Modems

Modem connections to the telephone service are made using two wires (*ring* and *tip*) that are used in a standard telephone jack. The wires are named for the plug wires used in the original telephone lines by which telephone operators would manually connect two telephones at the phone company switchboard. There are two versions of the telephone jack:

- **Half-duplex:** The RJ-11 has only two wires, which make up one line. Therefore, only one signal can be sent or received at a time.
- **Full-duplex:** The RJ-12 uses four wires to make up two lines; it can be used to simultaneously send and receive.

Modem Installation

With the advent of Plug and Play technology, Windows 98 and Windows 2000, and the USB port, installing a modem has become a simple process. Summaries of the general installation process for both internal and external modems follow.

Internal-Modem Expansion Card

As with installing any card or internal board, remember to take the proper precautions against electrostatic discharge (ESD), and, of course, back up your data before you open the computer case. Follow these steps:

1. Document: Check the current IRQ settings and I/O addresses in the computer. Make a note of available IRQs and addresses.
2. Configure IRQ and I/O settings for non_Plug and Play compliant systems: Set the modem to an unused COM port and IRQ.
3. Install the board: Physically install the board in an available expansion-bus slot.
4. Install any software: Follow the software setup routine and, if needed, fill in the modem settings and any dial-up connections the user requests for Internet access or for logging on to a remote system. To avoid generating any security concerns, do not ask for or accept account passwords. Show the user how to set that part of the connection personally.

   For older Windows 3.x or MS-DOS machines and non-Internet connections under Windows 95 and later versions, additional work might be required. Here is a sample of how this part of a typical Windows 3.x SYSTEM.INI file might look:

```
[386enh]
COM3Irq=5
COM3Base=03E8
```

5. Set up the command set: Any software that will access the modem must know the correct command set to use for that modem. This means identifying the type of modem so the software will use the correct AT commands. When all else fails, try using a Hayes-compatible modem.
6. Document your work: Write down *all* the new settings and changes.

External Modem

External modems are easier to install than internal modems because they do not run the risk of conflicting COM ports.

1. Connect to a COM port: Choose either COM1 or COM2. Be sure to confirm that the COM port you are using is assigned to the connector on the computer. If the computer is using a serial mouse, that will be using one of the COM ports, too, and this sets up a potential conflict. You can also use COM3 or COM4 if they are properly installed and configured.
2. Plug in the cabling: Connect the modem to its power source and to the computer. You will also have to connect a telephone line (RJ-11) from the wall jack to the modem.
3. Configure software: Configure the software to select the required COM port and the type of modem (command set) used by the specific modem installed.

Modem Speeds

When installing and using a modem, the primary factor to consider is speed. The multimedia World Wide Web requires far more speed than the simple data transfers of just a few years ago. Modem speed is measured in baud and bps.

Bits per Second

The actual modem speed, or rate, that data is transmitted, is measured in bits per second. If a modem modulates one bit for each baud cycle, then the modem speed is 2400 bps. If a 2400-baud modem modulates two bits for one cycle of time, the modem is said to have a speed of 4800 bps (not a baud rate of 4800). If four bits are modulated with one cycle of time, then a modem speed of 9600 bps is achieved.

Modem speed standards are given below:

CCIT Term	bps
V.21	300
V.22	1200
V.22bis	2400
V.23	1200 bps in one direction and 75 bps in the other
V.29	9600
V.32	4800 and 9600
V.32bis	14,400
V.32fast	28,800
V.34	28,800
V.42bis	38,400
V.90	56,000

NOTE

The term "bis" refers to the second revision of the standard.

Information Transfer Protocols

Communication relies on protocols. In order to ensure clear and clean communication without any errors, the device on each end must follow a very strict set of rules. If either device violates any of the rules, the communication will fail. These rules are called *File Transfer Protocols (FTPs)*.

All the necessary protocols should be included with the software that came with the modem. After communication is established with the host (usually the computer that receives the call), it can be asked what type of protocol to use. The call initiator can then select the matching protocol before starting a file transfer. *Both computers must use the same protocol*. There are five basic protocols used by modems.

ASCII

This protocol uses the standard ASCII character set, just like typing directly from a keyboard. ASCII protocol has no error-checking or compression features. It is simple, uncomplicated, and is used just to get simple character-based data. It is not a good protocol for transferring program files.

Xmodem

Xmodem is the next level of protocol. The advantage of Xmodem is that it includes error detection, which makes it more suitable for transferring program files. It transfers 128-byte blocks of data and one checksum (error-checking) character. The receiving computer calculates a new checksum and compares it to the one transmitted. If they are the same, the receiving computer transmits an ACK. If they are different, it sends back a NAK, and the transmitting computer then retransmits the data block. The protocol uses parity error checking, which is not perfect. If two errors were to occur—that is, if the first error were to change the parity bit, and the second error were to change it back to its original state—the second would cancel the first, and no error would be reported. The result can be a corrupted file or random characters on the display.

Ymodem

Ymodem is faster than Xmodem. Ymodem transfers data in 1024-byte blocks; therefore, less time is required to verify data with ACKs and NAKs.

Zmodem

Zmodem shares all the features found in Xmodem and Ymodem protocols. It also adds a few new features, including crash recovery, automatic downloading, and a *streaming* file transfer method. This is the protocol of choice for most situations.

Kermit

This protocol is rarely used today. It was the first of the synchronous protocols for uploading and downloading data to and from a mainframe computer.

Handshaking

Connections between a sending device (sometimes referred to as Data Communications equipment or DCE) and a receiving device (Data Terminal Equipment or DTE) are called *handshaking signals*. They ensure that each sending and receiving device is in sync with the other. The flow control of data between modems is handled by the modems themselves. However, the local flow control between modem and COM port can be set by the user. There are two types of flow control:

- **Hardware flow control:** This takes advantage of some of the extra wires in the serial connection between the modem and COM port. These wires are used to let the other device know that the DCE is ready to send or receive data. The wires are named RTS (Request to Send) and CTS (Clear to Send). Hardware handshaking is sometimes referred to as RTS/CTS.
- **Software flow control:** This uses special characters known as XON and XOFF to let the other device know that the DCE is starting to send data or that the data transmission is finished. Software handshaking is slower and not as dependable as hardware handshaking. Only some very old modems use software handshaking. If given a choice, always use hardware flow control.

Modem Standards

As with every other communication device, standards are needed to ensure that both sides "speak the same language." Modems have their own set of standardized communication conventions.

Error Detection

Some modems offer various forms of hardware error detection and correction. Such features usually require matching firmware in the modems at both ends.

Data Compression

Data compression is a means of shrinking files into smaller packets, which results in faster connections and requires less space on the host and client machine for storage.

Communication Standards

There were no standards during the early days of modem communication. The only way to ensure data transmission was to place identical modems at the sending and receiving ends of the transmission. Compatibility was a great concern, and proprietary modems were the norm.

Today, modems comply with several standards. There are two sources for these standards:

- Manufacturers have placed specifications of their modem functions in the public domain. These specifications can now be copied and used by any manufacturer. If enough manufacturers use a specification, it becomes a standard on its own merits.
- Standards committees are formed when there is enough interest expressed by users, vendors, or regulatory committees to develop a set of rules for a class of data or modems.

Modem-Speed Standards

The following table compares the standards set by the aforementioned organizations.

Standard	Baud	bps	Comment
Bell 103	**300**	**300**	U.S. and Canadian standard, now largely obsolete.
Bell 212A	**600**	**1200**	U.S. and Canadian standard.
V.21	**300**	**300**	Similar to B 103. However, B 103 modems are not compatible with V.21.
V.22	**600**	**1200**	Incompatible with 212A. This standard is primarily used outside the U.S.
V.22bis	**600**	**2400**	Offers world-wide compatibility.
V.23		**1200/75**	Provides split data transmission.
V.29		**9600**	Used for Class 3 fax machines.
V.32	**2400**	**4800/9600**	Provides error correction. Full-duplex.
V.32bis	**2400**	**14,400**	An improved version of V.32.
V.32fast	**2400**	**28,800**	An extension of V.32 and V.32 bis.
V.34 or V.fast	**2400**	**28,800**	Allows optional higher speeds of 31.2 Kbps and 33.6 Kbps.
V.90	**9600**	**56,000**	Allows optional higher speeds of 48.6 Kbps and 56.6 Kbps

56-K Modems

Telephone lines are capable of carrying 56 Kbps of data; however, conversion from analog to digital signals and back comes with a price. That price is a speed limit of 33.6 Kbps. Because many telephone systems are now digital, it is possible to transmit, in some instances, at a full 56 Kbps in one direction. The return data, however, is still limited to 33.6 Kbps. For these reasons, it is unlikely that you can achieve a full data transfer rate with a 56-K modem (56-K refers to a modem that transmits at 56 Kbps). In order to achieve the best performance, the following conditions must be met:

- Digital to analog conversion should be limited so that it takes place only once within the network. Each conversion slows the communication process.
- The host must be connected digitally.
- Both modems must support the 56-K technology.

There are currently three 56-K modem standards: K56flex, x2, and V.90. Unfortunately, these standards are not compatible at 56 K, so in order to achieve the highest possible speed, both modems must use the same standard. Several companies developed the K56flex standard, and U.S. Robotics developed the x2 standard. Currently, the V.90 standard is replacing both the K56flex and the x2, and most 56-K modems can be upgraded to this standard.

Troubleshooting

It can be very frustrating when a modem does not function as expected. However, follow these simple guidelines to determine whether the modem is really broken or something else is the culprit.

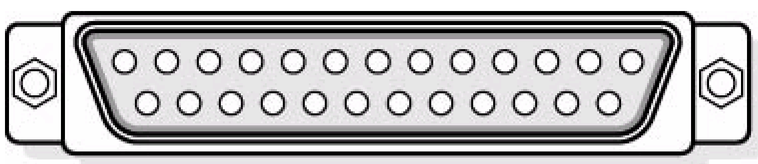
Possible Cause	Possible Solution
New hardware was added to the computer and now the modem doesn't work.	Check for IRQ and/or I/O conflicts.
New software was added to the computer and now the modem doesn't work.	Check for IRQ and/or I/O conflicts. Cards that were configured for software can be inadvertently changed by corrupted software or by the installation of new software.
The software says there is no modem.	Make sure the software is looking at the correct port. Reconfigure or reinstall the software (there might be a corrupt driver).
Modem works sporadically.	Try another modem type. Check the phone lines.
Modem does not hang up the phone line.	A power surge (lightning) can cause this problem. If manual disconnect and reconnect allows the modem to work, replace or repair the modem.

Cables and Connectors

Cables and connectors are critical to the operation of any computer peripherals. A computer professional must be able to identify and understand the various types of cables and connectors. This lesson discusses common connectors and their functions.

Parallel Printer Cables

Parallel printer ports and cables are used to connect printers and other add-on items such as CD-ROMs, tape drives, and scanners. Centronics Corporation invented the most common type. It is an 8-bit parallel connection with handshaking signals between the printer and the computer—these tell the computer when to start or stop sending data. A standard printer cable is configured with a 36-pin Centronics connector on the printer end and a standard 25-pin "D" (male) on the computer end. A standard 25-pin "D" (female) connector found on the back of the computer designates it as a parallel port.



Standard 25-pin D connector

The original parallel port was designed only to send information to printers and was unidirectional. However, some bidirectional communication was possible by manipulating the handshaking lines. Today, computer manufacturers have developed updated versions that allow better bidirectional communication while maintaining the original Centronics specification. The Institute of Electrical and Electronic Engineers (IEEE, pronounced "I triple E") developed a standard—IEEE 1284—to oversee the standardization of these ports.

IEEE 1284 Printer Modes

A vast array of printers is available, and in order to ensure that you are obtaining optimum performance, the printer, the printer driver, and the software using the printer must be configured for the same mode. The following table describes various printing modes and their capabilities.

Mode	Capabilities	Speed	Notes
Compatibility	8-bit output.	100-200-KB per	Original parallel

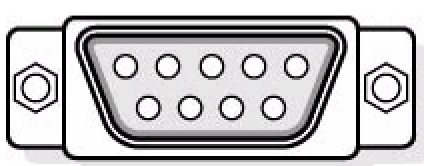
	Hardware handshaking. No DMA.	second out	port
4-bit	4-bit input using some of the printer's handshaking lines.	100-200 KB per second out; 40-60 KB per second in	HP Bi-Tronics mode
8-bit	8-bit I/O bidirectional.	80-300 KB per second	Original bidirectional port
ECP	8-bit I/O. Can use DMA.	>2 MB per second	Scanners and high-speed printers
EPP	8-bit I/O.	Up to 2 MB per second	Very flexible modes of operation

Parallel Pin Assignments

Just like modem cables, it is important for printer cables to have the correct pin connections. The following table describes the standard parallel pin assignments for the computer-end (25-pin) and the printer-end (Centronics) connectors.

Serial Port Cables

A serial port allows a computer to send data over long distances by converting parallel data to serial data. Typical computers will have one or two serial ports, usually designated as COM1 and COM2. The "standard" port is a 9-pin male connector on the computer, shown in the picture below. (There are also 25-pin cables available.)



Standard 9-pin D connector

The following table describes the pin connection for the 9-pin and 25-pin serial cable connectors.

9-pin	25-pin	Name	ID	Function
	1	Shield		
3	2	Transmitted Data	TX	Data sent from computer.
2	3	Receive Data	RX	Data sent to computer.
7	4	Request to Send	RTS	Computer is ready to send.
8	5	Clear to Send	CTS	"Other end" is ready to receive
6	6	Data Set Ready	DSR	"Other end" is ready to receive.
5	7	Signal Ground	SG	GND.
1	8	Data Carrier Detected	DCD	The modem has detected a signal from another modem.
4	20	Data Terminal Ready	DTR	Computer is ready to send.
9	22	Ring Indicator	RI	Modem detects line ringing.

Null-Modem Cables

Null-modem cables are used to directly connect two computers together without the need for a modem. The transmit and receive wires in the cable (wires 2 and 3) are switched to make the computers "think" they are using modems.

SCSI Cables

SCSI cables come in a variety of sizes depending on the type of SCSI used and the manufacturer of the device. Typically, internal cables are flat ribbon types and external cables are shielded bundles.

Keyboard Cables

Another peripheral device with a cable that we encounter and yet never think about is the keyboard. Because there are different types of keyboard cables (and connectors) and

because, on occasion, the technician might encounter a problem with a keyboard connector, they are worthy of mention.

Keyboards are manufactured in two different styles with different cables and connectors. Earlier versions used a 5-pin DIN connector (DIN stands for Deutsche Industrie Norm, the German national standards organization), and most new keyboards use a 6-pin DIN connector, the same as is used on a PS/2 mouse. Connectors are available to convert the 5-pin DIN to a 6-pin mini-DIN. Although they have a different number of pins, they use the same wires and pin outs. Data is sent serially to the keyboard using the keyboard interface. Data is written to the controller's input buffer to accomplish this. Keyboard data passes through pin 2 of the connector, while clocking signals move through pin 1. A keyboard reset, which can be connected to the system's reset line, is included at pin 3. Ground and +5-volt DC connections are applied to the keyboard through pins 4 and 5.

Identifying Cables and Connectors

Because printers and modems can both use 9-pin and 25-pin connectors and cables, a computer technician must be able to identify the function of the cables by their connectors. Other devices, such as monitors and game ports, use 15-pin connectors. Cable identification can be confusing, but it is important. The following table summarizes how to identify common cables and connectors.

Function	Computer Connector	Cable Connector
Communication (serial)	9-pin or 25-pin male	9-pin or 25-pin female
Printer (parallel)	25-pin female	25-pin male
Monitor (VGA and SVGA)	15-pin female (three rows of pins)	15-pin male (three rows of pins)
Monitor (MGA and CGA)	9-pin female	9-pin male
Game port (joystick)	15-pin female (two rows of pins)	15-pin male (two rows of pins)
Keyboard	5-pin DIN female or 6-pin DIN female (PS/2)	5-pin DIN male or 6-pin DIN male (PS/2)

Troubleshooting Cables

Cables and connectors are a very common source of problems. Here are a few suggestions for troubleshooting:

- If a peripheral device doesn't work, always check the cables, especially if the device has been working recently.
- Always check for loose connections.
- Check for bent or broken pins on the connector. Bent pins can sometimes be repaired; however, they will always be susceptible to damage later, because the pin has been weakened. It is a good idea to mark these connectors and use them with care. A better idea is to replace them.
- If a connector or cable doesn't fit or if you have to push hard to make the connection, something is wrong. Either a connector has been damaged or it is not the right match.
- Check for worn or frayed cables. Replace if necessary.
- Make sure you have the right cable. Some, such as null-modem cables, look just like standard communication cables, but will not work with a modem.
- Always be wary of "homemade" cables.

Chapter 13

Basic Electrical Energy

OBJECTIVES

Power

Electrostatic Discharge (ESD)

Safety and Electrical Power

Power

Computers run on electrical energy. Without it, a computer might as well be a paperweight. Because every component of a computer needs power to run (whether plugged into a wall outlet or a battery), a computer professional must understand the basic principles that govern electricity and electrical energy. This chapter introduces you to these principles.

The energy or power to drive a computer is derived from electricity. Whether it uses 110 volts alternating current (AC), the U.S. standard, 220 volts AC, the European standard, or direct current (DC) from a battery, a computer is useless without a steady, reliable source of power. When we encounter problems with a computer, it is crucial for us to be able to test the entire power system. This lesson covers the basics of power and electricity.

Understanding Electricity and Electrical Energy

What is electricity? The meaning of the word varies with the user. Electricity to physicists is the primal property of nature, and they call the power delivered at the wall socket and stored in batteries electrical energy. Most people, including computer technicians, are less fussy, often using the term electricity to refer to both:

- The form of energy associated with moving electrons and protons.
- The energy made available by the flow of electric charge through a conductor.

For our discussion, we mostly talk about the flow of energy used to run computers electrical energy and not worry about the fine points of scientific philosophy.

Some Definitions

For our discussion we employ the following definitions:

- Electricity: The form of energy associated with charged particles, usually electrons.
- Electric Charge: When charged particles move in tandem, fields are generated, producing energy.
- Electrical Circuit: The path taken by an electrical charge.
- Electric Current: When an electric charge is carried, or flows through a conductor (like wires), it is known as a current. A current-carrying wire is a form of electromagnet. Electric current is also known as electron flow.

- Power: The rate at which an amount of energy is used to accomplish work. Electrical power is measured in watts, which is determined by multiplying voltage by current.
- Conductors: Materials that can carry an electrical current. Most conductors are metals.
- Resistance: A quality of some materials that allows them to slow the speed of an electrical current, producing heat, and sometimes light, in the process.
- Insulators: Materials that prevent or retard the electrical current of electrons.
- Ampere: A measurement of current strength, equal to 1 Coulomb per second. Coulomb's Law: two charges will exert equal and opposite forces on each other. Opposite charges attract and like charges repel.
- Ohm: A unit of electrical resistance. Ohm's Law states that voltage is equal to the product of the current times the resistance, or Voltage = Current x Resistance.
- Volt: The unit of electromotive force, or potential energy, that when steadily applied against a resistance of 1 ohm, produces a current of 1 ampere.
- Voltage: The potential energy of a circuit.

Ohm's LAW

In addition to defining terms, we need to understand some basic principles that are applied in testing electrical devices. One formula that all computer professionals should know is *Ohm's Law*. From this formula, or a derivation of this formula, all basic power calculations can be performed.

Ohm's Law states that the current (electrons) flowing through a conductor, or resistance, is linearly proportional to the applied potential difference (volts). A conductor is any medium, usually metal, that allows the flow of electrical current. Resistance is any device or media that resists the flow of electrons. In mathematical terms, this means:

Resistance: $R=V/I$

Current: $I=V/R$

Volts: $V=IR$

In these formulas, R = Resistance in Ohms (W), V = Voltage, and I = Current in Amperes.

NOTE

In some cases you may find E used instead of V in formulas expressing Ohm's Law. The E stands for "electromotive force," and is often used by engineers as a more precise technical term for their measurements.

By memorizing any one of these formulas, the other two can be easily derived using simple algebra. For example, the voltage (potential energy of the circuit) is equal to the amperage (the current or flow of electricity) multiplied by any resistance to that flow of electricity. The more resistance there is in a circuit, the lower the current flow for a given voltage.

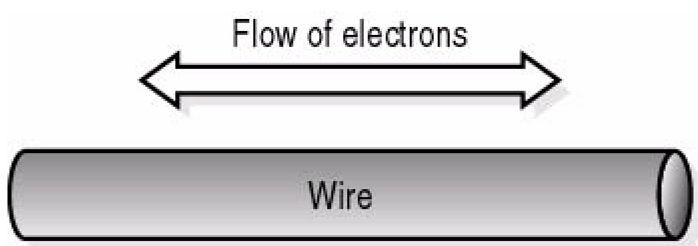
PCs and Electrical Power

That PCs use electrical power to operate is no surprise, even to the casual user. The technician must understand the different types of electrical energy and how they work inside the PC. A PC's electrical power can come from a wall outlet, in the form of alternating current, or from a battery in the form of direct current.

AC Power (Alternating Current)

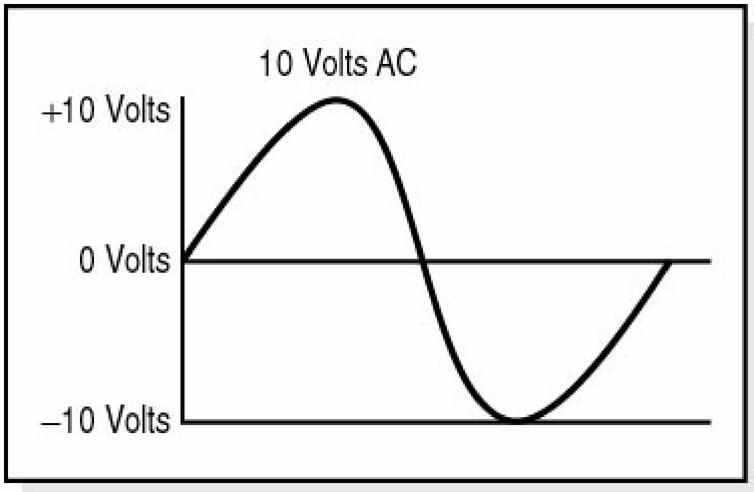
AC power is what most people think of as electricity. It comes from the wall, and powers most of our lights and household appliances.

AC power is man-made, using generators. As the wire coil inside the generator rotates, it passes by each pole of unit magnet(s) producing an electric current. When it passes the opposite pole, the current reverses, or *alternates*, the direction of flow (see Figure 13.1). The number of revolutions made by the generator per minute is called its frequency. In the United States, power companies run their systems at 60 turns per second to produce a high-voltage, 60Hz (cycles per second) alternating current as they rotate. The power system drops the voltage in stages before it is connected to the consumer's home or business.



Flow of electrons

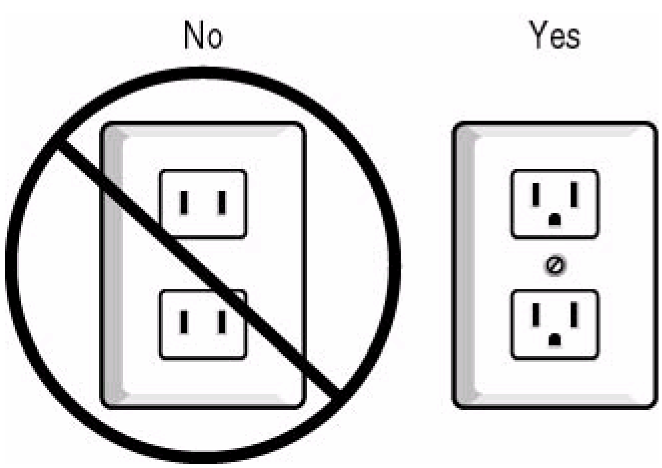
The power company delivers AC power to our homes or businesses with three wires. Two of the wires are *hot*, meaning that they carry a charge. One, the bare wire that runs from the breaker box to the power pole, is *neutral*. The measured voltage between the two hot wires is between 220 and 240 volts AC (VAC), while the measured voltage between either of the hot wires and the neutral wire is between 110 and 120 VAC. These voltages, which are called *nominal* voltages, can vary by plus or minus (±) 10 percent.



AC volts

Typical electrical outlets are connected between one of the hot wires and the neutral wire. These outlets are usually three-prong connections. The smaller rectangular hole is the hot, the larger rectangular hole is the neutral, and the small round hole is called the *ground*. The ground wire is used as a safety wire. In the event of a short circuit, a large flow of current (amps) is discharged all at one time. This short, high flow of current will burn out circuits unless it can be safely sent somewhere else. Electricity will always seek the path of least resistance to ground. By providing this wire, a short circuit will cause less damage by providing a path for safe dissipation of the current. To provide a safe working environment for the computer and yourself, make sure that this wire is properly installed.

Older structures might have two-wire electrical outlets without the ground wire. An electrical outlet without grounded plugs and the third ground wire is unacceptable for use with a computer. An extension cord without a ground wire is also unacceptable.



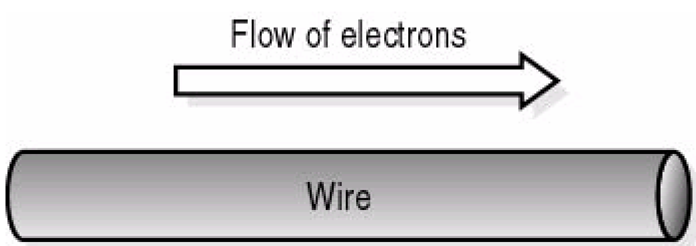
The proper type of outlet includes a ground

CAUTION

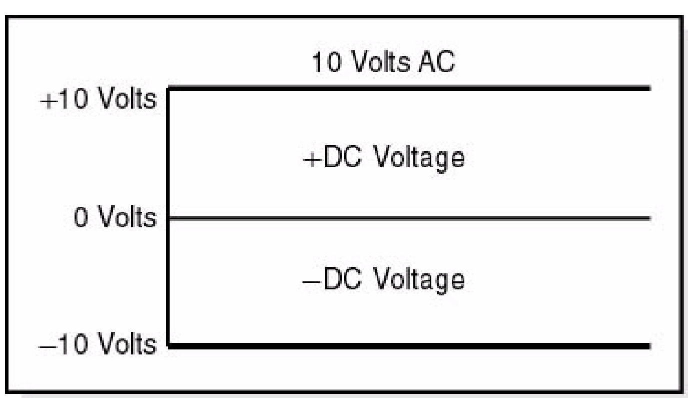
A short circuit can cause physical damage to equipment and personnel. It can cause a fire, component damage, permanent disability, or even death. The ground plug provides a direct connection to ground, giving the electricity an alternate path away from equipment and people.

DC Power (Direct Current)

Alternating current is used for transporting low-cost power to end users. But a computer's electronic components won't run on AC power—they need a steady stream of direct current. The PC's power supply performs several tasks, but the main function is to convert AC into DC. A computer's power supply combines two components to handle this job: a *step-down transformer* and an *AC/DC converter*. The AC adapters used for laptop computers, many low-cost ink-jet printers, and many other consumer electronics do the same thing—turn alternating current into lower-voltage direct current. As we have seen, DC is electrical energy that travels in a single direction within a circuit. (The electrical energy in a thunderstorm is another example, but not very practical in electronic applications.) DC current flows from one pole to another, hence it is said to have *polarity* (see Figure 13.4). The polarity indicates the direction of the flow of the current and is signified by the "+" and "-" signs.



DC power



DC voltage

Measuring Electricity

A computer professional should know how to use a multimeter—sometimes called a VOM (Volt-Ohm Meter) or a DVOM (Digital Volt-Ohm Meter). An electrical test meter is probably the best (and most practical) tool for troubleshooting electrical problems. It is not necessary to be an "electronic technician" to use this tool effectively.

The Multimeter

A multimeter is an instrument that is used to measure several aspects of electricity. All multimeters are designed to provide at least four major measurements:

- AC voltage
- DC voltage
- Continuity
- Resistance

A multimeter consists of two probes, an analog or digital meter, and a multiposition switch to select the type of test you wish to perform.

Testing AC Power

On any new building installation, failure to properly test AC outlets can result in damaged or destroyed equipment, as well as possible injury and electrocution. In the event a wiring error was made that causes the voltage to be outside of the specifications (either two high or too low), problems are sure to arise. Don't take for granted that the building power supply provides the correct voltage, or that all of the other inputs are wired correctly.

When testing an AC power source, check these three things:

- Is the *hot* wire sending the correct voltage, and is it wired to the correct pin?
- Is the *neutral* wire connected to ground and to the correct pin?
- Is the *ground* wire connected to ground and to the correct pin?

Testing AC Outlets with a Multimeter

The first step when testing an AC outlet is to set up the multimeter. Then you need to know how to read the meter. You can also use special equipment if the multimeter does not provide enough information.

Setting Up the Meter

Basic multimeter usage with AC circuits is quite straightforward:

1. Attach the black test lead to the negative (-) marked hole. In some low-cost meters, leads are permanently attached to the meter.
2. Attach the red test lead to the volts (+) hole. Be careful—if this lead is placed in the wrong hole (ohm or amp) it can cause permanent damage to the meter.
3. Set the selector switch to AC volts (this choice is often denoted by red lettering). If there are multiple selections, use the highest setting possible (if voltage is unknown), or select one level higher than the estimated voltage. For standard household outlets, 200 VAC is a good selection. Some digital meters use "Auto-range" and don't need any selection except AC volts.

Reading the Meter

After the meter is set up, you are ready to test a wall outlet. There are three tests to perform. With AC voltage, it does not matter which lead is placed in which connector:

- **Hot to Neutral:** Place one lead in hot (smaller of the two vertical slots) and the other in neutral (larger of the vertical slots). The reading should be between 220 and 240 volts AC.
- **Hot to Ground:** Place one lead in hot (smaller of the two vertical slots) and the other in ground. The reading should be between 220 and 240 volts AC.
- **Neutral to Ground:** Place one lead in neutral (larger of the two vertical slots) and the other in ground. The reading should be 0 volts.

Using AC Testers

An alternate method for testing electrical outlets is to purchase an AC tester. These small devices are made especially for testing outlets and can be purchased at any home improvement or electronics outlet store. By simply inserting the tester into an outlet, all voltages for all combinations can be tested at the same time. Many testers provide several LEDs that tell whether or not each function passes the test. This device is not as accurate as a multimeter but it is more convenient. It will provide a pass/fail indication rather than an accurate voltage reading.

Testing AC Ripple

The function of the power supply is to convert AC to DC voltage. When working properly, a pure DC signal will be produced. However, sometimes, as the power supply ages, its ability to produce pure DC falters. A power supply uses electrolytic capacitors (discussed later in this lesson) to filter or smooth the voltage after it has been converted from AC to DC. These capacitors are second only to fuses as the part of a power supply most likely to fail. When an electrolytic capacitor begins to fail, it allows more and more

AC voltage to pass through. This small amount of AC voltage is superimposed on top of the DC voltage and called *noise* or *ripple*. To test for ripple, set a meter to read AC. Then connect a .1mfd (microfarad) capacitor to the red lead. With the power turned on, measure the DC voltage. Any ripple present will be displayed as AC voltage.

Testing Resistance

Resistance is an opposition to the flow of current through a conductor. Resistance is measured in *ohms*. The symbol for an ohm is W. Resistance is measured by placing one lead of the meter on each side of the circuit or component to be measured. Taking resistance measurements for a component while it is still soldered in its circuit can lead to inaccurate readings because any other component connected to the circuit can affect the total resistance measured. Unlike voltage checks, you should test resistance with the power off. If a meter is set up to read resistance, you will damage it if you connect it to an electrical outlet.

NOTE

Be careful when measuring resistance. If the meter is set too high or the resistance is too high for the meter, you will get an inaccurate reading. Also, before taking a measurement, be sure that any charge stored in a capacitor is properly discharged. Refer to the applicable product manual for details.

Testing Continuity

Continuity is a term used to indicate whether or not a connection exists between one point in a circuit and another. It is used to determine the presence of breaks in wires and electrical circuits. If no continuity setting is available, use the resistance setting (see the next section). If the multimeter measures infinite resistance, there is no continuity. This indicates a break in the line. If the multimeter shows little or no resistance, there is continuity and the circuit is complete.

Testing DC Voltage

Testing for DC voltage is the same as testing for AC voltage, but with one important difference: DC voltage is sensitive to polarity. As mentioned earlier, DC voltage has a positive pole (+) and a negative pole (-). When measuring DC voltage, it is important to place the positive (red) lead on the positive side and the negative (black) lead on the negative side of the circuit. If the leads are positioned backwards, the polarity of the reading will be the opposite of what it should be.

CAUTION

When using an analog meter (one with a dial and needle), connecting the leads backward will cause the needle to move in the opposite direction, possibly damaging the meter.

Testing a Power Supply

Find out if the client is having any problems with flickering lights, intermittent problems with other appliances, or is using a power strip with too many connections for the rated use; improper loading of the circuit, not the PC itself, can be the problem.

A bad power supply can cause intermittent lockups and unexpected computer reboots. Erratic problems encountered during booting and changed or erased CMOS information can also be traced to a failing power supply. Bad power supplies have been known to destroy data on mass-storage devices. There are two types of tests for power supplies: a basic test used to verify voltages and an advanced test for checking its internal components.

Basic Voltage Test

The only purpose of this test is to verify the existence and value of voltages. With time, most power supplies show their age by a reduction in voltage. This voltage drop will show itself in both the 5-volt and the 12-volt outputs, but is more pronounced on the 12-volt side.

Prepare the Meter for the Test

Again, meter preparation is quite simple:

1. Connect the black lead to the common (-) connector and the red lead to the voltage (+) connector.
2. Turn the test selector to DC volts. If the meter has an AC/DC switch, be sure it is set to DC. If the meter does "auto range," set the range to 15-to-20 volts.

Testing the Voltages

The best place to check voltage is at the power supply's P8/P9 or ATX power connectors (see Chapter 5, "Supplying Power to a Computer"). For P8/P9 systems use the instructions below:

1. Place the meter's black (ground) lead on the black wire connection and its red (positive) lead on the yellow (+12 volt) connection.
2. Record the voltages. A good power supply will provide a voltage between 11 and 13 volts DC.
3. Replace the power supply if the voltage reading is less than 10.

Be sure to reverse the leads when using an analog meter to check negative voltages. This is not necessary with a digital meter because it will simply show a negative sign with the reading.

When No Voltage Is Present

If you have completed the basic voltage test and no voltage is present, the problem may not be the power supply. It might, instead, be caused by an excessive load on the system due to another piece of hardware. To determine if that is the case, try the following procedure.

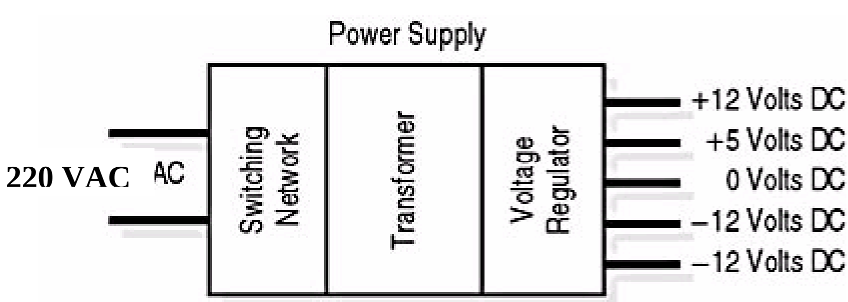
Isolating the Problem

First you should test the hardware:

1. Disconnect the Molex leads from the power supply.
2. Connect the meter leads as described in the previous sections.
3. Turn off the AC power.
4. Disconnect all the Molex plugs from the devices.
5. Turn the power back on. If power is present at the motherboard, one of the devices is bad and is causing a drain on the power supply.
6. Reconnect each Molex plug, one at a time, and test the power. When the power drops out, you have located the offending device.

Advanced Testing

The basic test is designed to quickly isolate the power supply as a problem. In most cases, if the test proves the power supply to be defective, it may be more cost-effective to replace the power supply than to try to repair it. Advanced testing requires a working knowledge of power supplies and removal of the power supply and its cover. There are three sections to a power supply: the *switching network*, the *transformer*, and the *voltage regulator*.

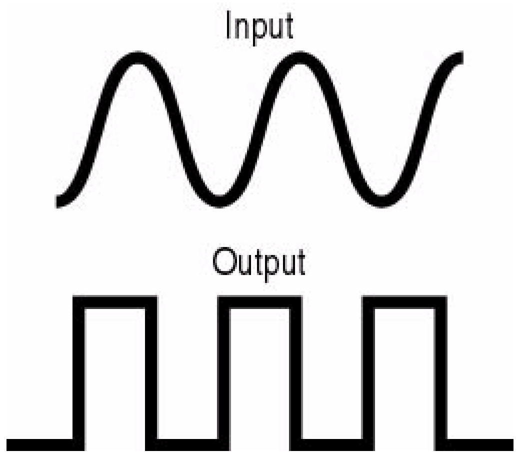


Power supply

The Switching Network

AC power coming from the power company is imperfect. It is not uncommon to have sudden increases in voltage called *spikes* or decreases in voltage called *sags*.

To smooth out the power sent to the electronic components, a PC has basic line conditioning capability—a switching network. The better the power supply, the more sophisticated the network. The main components found in a switching network are a fuse, capacitors, rectifiers, and switching transistors.



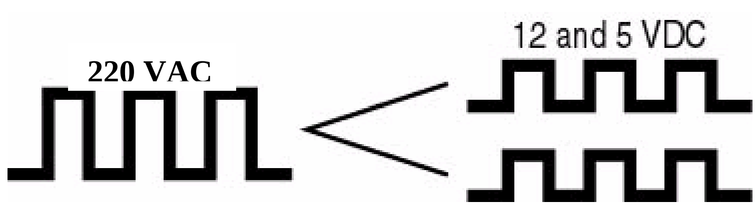
Switching network

The switching network performs the following three tasks:

- Filters electrical noise (spikes, sags).
- Eliminates frequency changes (ensuring that current stays at a constant 60 cycles).
- Converts AC *sine wave* signals to AC *square wave* signals.

The Transformer

The transformer reduces the voltage of the *square wave* DC into separate 12-volt and 5-volt square wave AC circuits.



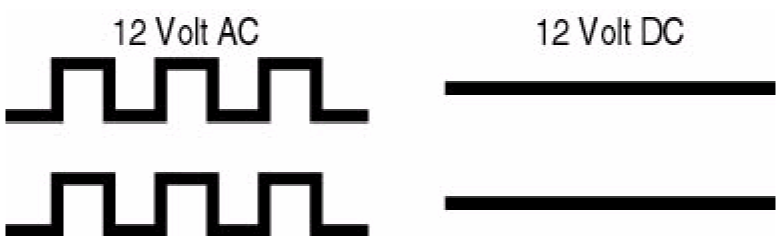
Transformer voltage

The Voltage Regulator

The voltage regulator receives the low-voltage AC outputs of the transformer and converts them to clean DC power. The main components in this section are rectifiers, capacitors, and coils.

The voltage regulator section performs three functions:

- It uses rectifiers (diodes) to convert the square wave AC output of the transformer into DC output.
- It regulates the voltage to a constant output level and uses capacitors to remove any ripples that are present.
- It monitors the amount of current used by the computer circuits and adjusts the switching network using a special circuit called the feedback circuit. This compensates for variations in load on the power supply.



Regulator section voltage

CAUTION

Do not open the power supply while it is plugged in, and do not open the power supply until it has been discharged. The power supply can carry dangerous levels of power even when disconnected. Only a properly trained technician should ever open a PC power supply. Given the cost of a power supply, there is no good reason to disassemble one; defective units should be replaced.

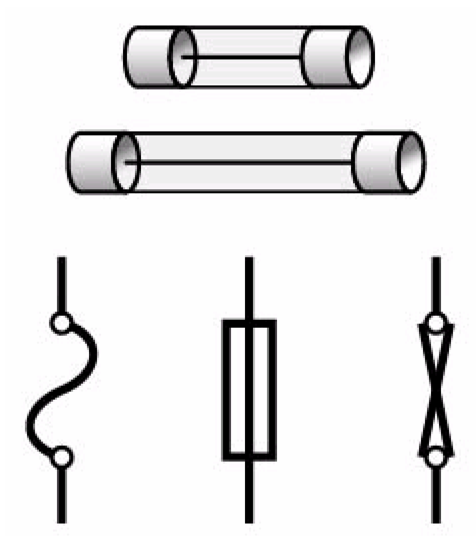
Electronic Components

As a computer professional, you should be familiar with the more common types of electronic components within a power supply. Here is a description of the basic components found on circuit boards inside a computer.

Fuse

Before the advent of the circuit breaker, fuses were common in the home and office. A fuse serves one purpose—to fail—and thus cut the flow of power in the event of a current

load that has exceeded the safe capacity of the system components to absorb. While fuses come in many shapes and sizes, a PC fuse is almost always a small, clear glass tube with metal caps on each end and a wire inside the tube to electrically connect the two caps. In general, the thicker the wire, the more current it can conduct before failing. When a fuse fails, the wire will melt or be broken. You can check for a "blown" fuse by seeing if the wire is intact or broken. The amperage (A) rating (stamped on the metal cap) indicates the maximum current the wire is rated to conduct. Be sure not to exceed to the rated limits of the PC design for a fuse, because an excess power load can damage or destroy the system.

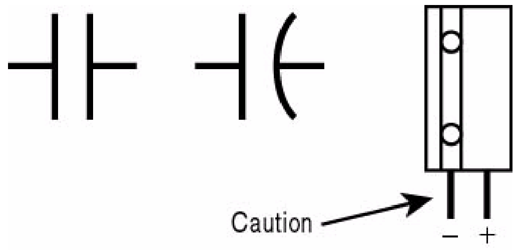


Fuses

If a fuse in a specific location fails more than once or repeatedly, the system is being overloaded, and you need to isolate the problem causing the failure. Fuses are often found on power supplies and many external components. If a fuse fails, try first replacing it with another of the same rating. If the replacement also fails, the fault probably lies with the motherboard or another internal part.

Capacitors

A capacitor is an electrical component used to hold an electrical charge. In photography, electronic flashes use capacitors to build up power before a picture is taken and to vary the amount of power used in a flash to control the exposure. In PCs, they are often used to regulate the flow of current to areas of the system circuits for a short period of time. Some are fixed-capacity models, whereas others can absorb or hold variable amounts of power. The amount of electrical current a capacitor can control is called capacitance, measured in microfarads.

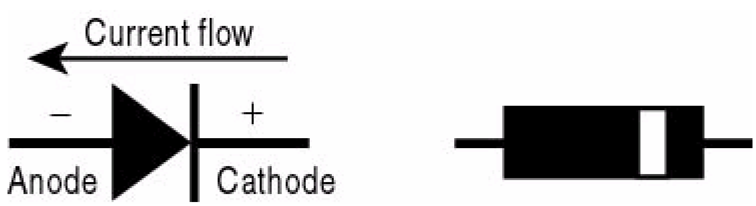


Capacitors

Most PC power supplies employ an *electrolytic* capacitor. These devices are able to retain a significant charge for long periods. You should work with such components only if you are properly trained to know how to release any residual charge before disconnecting, testing, removing, or replacing one. Failure to follow safe procedures can result in injury or death to you, and damage to the system. These capacitors have a distinct polarity (negative and positive) to their two leads.

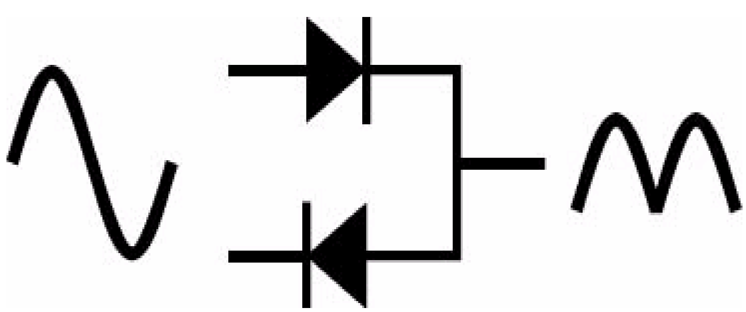
Rectifiers and Diodes

Rectifiers are devices that convert AC power into a DC form (rectification). A diode is a device that lets current flow in only one direction. Two or more diodes connected to an AC supply will convert the AC voltage to DC voltage.



Diodes

Single diodes are generally used to convert AC current to pulsating DC current. Two diodes working in parallel produce half-wave rectification, resulting in a pulsating direct current. Four diodes produce full-wave rectification, with a continuous stream of pulses.

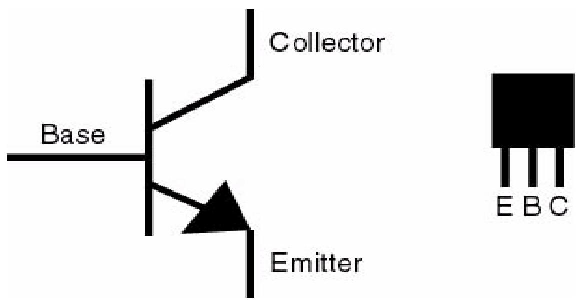


Half-wave rectifier

Normally, a computer technician does not test at this level; however, diodes can be tested with a multimeter. With the power turned off, test for resistance across both leads of the diode. Then reverse the leads of the multimeter and test again. A good diode will exhibit low resistance in one direction and high resistance in the other.

Transistors

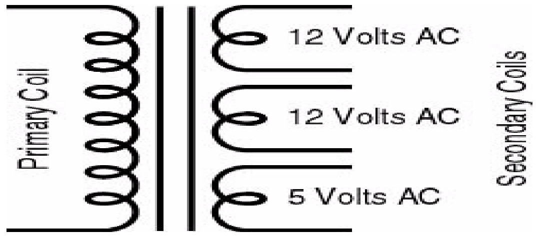
The invention of the compact, power-efficient, and reliable transistor created the modern electronics industry; replacing bulky, power-hungry, and temperamental vacuum tubes. Transistors are basically a pair of diodes connected in series with an "on-off" switch. Varying the voltage sent to a transistor turns the switch on or off. Early computers used vacuum tubes as switches and were so large that technicians could actually step inside the larger ones to plug in or remove the tubes in order to program them.



Transistors

Transformers

The most common forms of electrical transformers are step-down or step-up devices. A step-down transformer decreases the transformer's voltage on the output side; a step-up model increases it. Both have a primary wire coil connected to other coils secondary coils joining two or more AC circuits. Electronic transformers generally contain stacks of thin metal-alloy sheets, known as laminations, with coils of copper wire wound around them. They are commonly employed in a circuit along with a rectifier which, as we have seen, supplies DC power to the equipment. In the PC power supply, the transformer's secondary coils are used to provide 12-volt, 5-volt, and 3.3-volt outputs used by various components



Transformer

Testing a Transformer

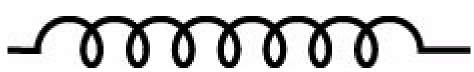
As noted, a transformer is made up of several coils of wire. Because each coil in the transformer is continuous, each can also be tested for continuity. Follow these steps:

1. Disconnect the power.
2. Discharge all capacitors.
3. Check the bottom of the circuit board of the power supply to be sure all leads have been disconnected. The primary connections, as well as the secondary connections, can be located below the transformer.
4. Configure the multimeter to measure continuity (or resistance).
5. Simultaneously touch each lead of a multimeter to one of the pairs of contacts.

A good transformer will show a reading of low resistance. A very high reading could indicate that one of the coils is broken.

Inductors (Coils)

Inductors, commonly called coils because of their shape, are loops of conductive wire. Current passing through the inductor sets up a magnetic field. This field reduces any rapid change in current intensity. Inductors can also be used to distinguish between rapidly and slowly changing signals in a circuit.



Coil

Testing a Coil

Since an inductor is simply a wire coil, it can be tested for continuity in much the same way as a transformer is tested. Visually inspect the wire for deterioration. If it shows signs of breakage or burned areas, it should be replaced. If the wire looks good, follow up with a conductivity test. Turn the system power off. Disconnect one lead to the coil (this might require a soldering iron) and connect one meter lead to each end of the coil. A null or low reading indicates continuity. A reading of high or infinite resistance indicates a lack of continuity. Replace the coil.

Electrostatic Discharge (ESD)

Causes of Electrostatic Discharge

The human body has an electric field and under the right (and very easy to obtain) conditions can generate a tremendous amount of voltage, often referred to as static electricity.

ESD occurs when an imbalance in the amounts of positive and negative electrical charges on the surface of an object is released. The most dramatic example of ESD is lightning, which splits trees as easily as it lights up the sky. The amount of energy released when you touch a metal object can be quite large. The buildup of energy with nylon clothes can easily reach 21,000 volts. About 750 volts are required to produce a visible spark with ESD, while a mere 10 volts or so can ruin a computer chip.

The actual amount of energy in a given ESD event depends on the types of materials involved (wool fabrics generate less than nylon), the humidity (low humidity offers less resistant to the discharge), the amount of physical energy (friction) involved, and how quickly the energy is released.

NOTE

ESD does not have to be seen (a spark) to do damage to electronic components. Voltages lower than 10 volts can damage some parts.

ESD Damage

Over the years, engineers have produced smaller and smaller components, which operate at lower and lower voltages. Their goal is to reduce size, cost, and operational heat production. Those are worthy goals, but because of reduced component size, they present a smaller target with less resistance to power surges. This makes the parts more susceptible to damage from ESD.

The amount of damage and resulting problems caused by ESD can be divided into three categories:

- **Catastrophic Failure:** This is sometimes referred to as "frying" or "smoking" a part because of the heat (and sometimes the noise and smoke) generated during the failure. Mishandling and misapplication of a power source, cable, or test instrument are the most likely causes. Care in opening, installing, cabling, and testing are the best ways to prevent this type of ESD damage.

- **Upset Failure:** An ESD can produce an erratic fault in a component. This kind of problem is very difficult to detect and repair, because the failure is intermittent. It is easy to blame the operating system or a program operation for the data loss or system crash. The best way to diagnose and correct this type of problem is to remove and replace suspected parts until the failure stops appearing.
- **Latent Failure:** This type of failure weakens the actual transistor. The part will seem normal in most operations and will frequently pass quality control and conformance tests. Like upset failures, these can be very difficult to isolate.

Preventing ESD

Prevention is the best defense against ESD, and the first step of prevention is to understand the source. The number one cause of ESD damage is improper handling of electronic devices. A semiconductor device can be damaged by ESD during handling before it is installed.

The key to ESD prevention is to keep all electronic components and yourself at a common electrical potential. This usually means ground potential, or zero volts. Maintain a habit of "grounding" yourself to the computer chassis whenever you attempt a repair. There are times when it is not practical, or convenient, to wear a ground strap. At such times, touching a part of the metal chassis before removing devices will bring you and the computer chassis to a common voltage. Don't move around while installing or handling a part; doing so can generate additional voltages, negating any effort you have made to eliminate ESD.

All repair shops and workbenches should have proper ESD-suppression devices, and technicians should use them whenever working with exposed parts. These devices include:

- **Antistatic mats:** Nonconducting pads placed on the work surface and on the floor in front of the work area.
- **Antistatic wristband:** A wristband with a grounding strap connected to the chassis of the PC.
- **Antistatic pouches:** A sealed, antistatic pouch used to store any sensitive electronic device, including hard disk drives, when they are not installed in a computer.
- **Antistatic pad:** An insulating foam pad in which individual chips with exposed pins should be embedded when they are not installed in a computer.

Safety and Electrical Power

Computers are electronic equipment and therefore consume electricity. Although most components in a computer operate on low voltages in the safe range between 3.3 and 12 volts DC the main power source and the monitor use high voltages. This lesson sets out some basic high-voltage electrical safety guidelines.

Electrical Safety Is Your Responsibility

Standard wall outlets provide a nominal 220 volts AC and are rated to deliver currents between 7.5 and 10 amps. Under certain conditions, it is possible to receive a lethal shock from much lower voltages than these. Inside a computer, and especially the monitor, voltages as high as 30,000 volts can exist even *after* the power is turned off.

It is vital to follow basic electrical-safety guidelines when servicing a computer. There is no substitute for good old common sense. However, here are a few tips:

- When in doubt about the correct way to safely service a part of a computer, don't do it. Have an experienced professional do the necessary work.
- Always use grounded outlets and power cords.
- Switch the power off and disconnect all equipment from its power source before removing any covers.
- Always replace blown fuses with fuses of the correct rating and type.
- Do not work alone you might need help in an emergency.
- Remove all jewelry and any wristwatch. These are conductors and can cause short circuits.
- Have trained personnel service computer power supplies and monitors; these devices use and store potentially lethal voltages (often for days or longer).
- Work with one hand. Using two hands can cause a direct circuit, via your heart, from one object to another.

In the U.S., common AC wiring uses the following color coding:

Connection	Color
Live or hot	Black
Neutral	White
Ground	Green or bare copper

Chapter 14

Upgrading a Computer

OBJECTIVES

Computer Disassembly & Reassembly

Upgrading Computers

Computer Disassembly and Reassembly

Computer upgrades are among the most common tasks performed by a computer technician. With new technology coming out every day, it is a constant struggle to stay up to date. Before you can upgrade or repair a computer, you must know how to take it apart. A thorough understanding of how to disassemble a computer and put it back together is required by any computer professional. This chapter provides guidelines for successfully disassembling, upgrading, and reassembling a computer.

Documentation to Collect Before Starting the Job

The following list provides examples of the types of documentation you should assemble before you begin a repair.

- A computer configuration sheet. A sample computer configuration sheet is shown in the picture below.
- Copies of the computer and/or motherboard documentation.
- A list of all installed expansion cards. If possible, include the date on which they were originally installed.
- Copies of the operating-system documentation (especially if you are not familiar with the system).
- A plan of action. Writing down a checklist of tasks and related tools and parts *before* starting a project can help you keep focused and stay on target. Remember, plans can always be changed; but without a plan, you could find yourself wandering aimlessly through the project and perhaps getting sidetracked or lost.

Questions to Ask Yourself Before Starting the Job

Carefully consider the following questions before you open the case of any computer:

- Is this the right computer?
- Why am I taking it apart?
- Do I have everything necessary to do the job?
- Do I need more information before starting this job?
- Might any components of this machine be proprietary hardware? If so, do I have the right tools and parts to do the job?
- Do any of these tasks require the assistance of a third party—for example, monitor adjustment?

Recommended Hardware

A computer professional does not need a large toolbox; only a few basic hand tools, a handful of floppy disks are required to solve most computer problems. Most PCs can be opened and most parts removed and replaced with a pair of screwdrivers. Be careful when working on a proprietary machine, however—special tools are often required. A small canvas bag or a briefcase will generally be sufficient to carry everything you need.

The following table lists and describes the hand tools that will meet most needs.

Tool	Description
Screwdrivers	Two of each (one large and one small) flathead (regular) and Phillips (sometimes called a cross) screwdrivers are usually sufficient. Avoid magnetic screwdrivers: although they are convenient for picking up lost screws, their magnetism can cause problems.
Torx driver	Used to remove the odd star-shaped screws found on some proprietary computers. Sizes T-10 and T-15 should meet the needs of most computers.
Tweezers	Very convenient for picking up small parts (for instance,screws). You might consider the long plastic variety; these don't conduct electricity and hence won't create any short circuits.
Needlenose pliers	Can be used to pick up dropped items and to hold or loosen screws, nuts, and bolts.
Chip removers	Although optional, these are very useful when changing video RAM or other (older) RAM chips that are pushed into a socket.
Tube for small parts	A short plastic tube (with caps on both ends) will keep loose screws and small parts from wandering.
Compressed air	A can of compressed air is helpful to remove dust.
ESD tools	An antistatic wristband is a must. Antistatic mats and antistatic bags are also helpful.
Multimeter	A small, digital meter that is capable of measuring volts (AC and DC) and ohms (resistance or continuity) is all that is needed.

Flashlight	A small (bright) light for illuminating those hard-to-get-at places.
Nut driver set	Sizes 3/16-inch, 7/32-inch, and 1/4-inch.
Hemostats	Good for picking up and holding small parts. Straight hemostats will work most of the time. However, curved ones will get into those small places that the straight ones can't reach.
POST Card	A POST card can be used to see what the error messages during system start are being sent when no data is being sent to the display.

Recommended Software

Don't feel compelled to carry an entire arsenal of arcane software. A small number of commonly used programs can serve most of your needs.

Bootable Floppy Disk

You'll want to compile and carry bootable floppy disks for each operating system that you encounter. These should contain the following files:

• ATTRIB.EXE	• **FORMAT.EXE (.COM)**	• QBASIC.EXE
• **COMMAND.COM**	• HIMEM.SYS	• SCANDISK.EXE
• DEFRAG.EXE	• LABEL.COM	• SHARE.EXE
• EDIT.COM	• MEM.EXE	• SIZER.EXE
• EMM386.EXE	• MEMMAKER.EXE	• SMARTDRV.EXE
• EXPAND.COM	• **MSCDEX.EXE**	• **SYS.COM**
• **FDISK.EXE (.COM)**	• **MSD.EXE (.COM)**	

NOTE

These files will just barely fit on one 3.5-inch high-density floppy disk. Files listed in **bold** are essential.

A Windows 98 startup disk is also a good item to carry. This is a bootable disk that will also load all drivers needed to run a CD-ROM on most PCs.

TIP

The utility MSD.EXE is a good diagnostic tool that can determine which hardware options are installed on a computer system without the need for you to remove the case. MSD.EXE is also a great tool for diagnosing software conflicts.

Operating-System Disk

Make sure copies of the original operating-system disk (or CD) are available. If it becomes necessary to install one or more components that were left out during the original installation, the computer might require verification of serial numbers (the original disk #1) before any additional files can be installed. Windows 95 and 98 and Windows NT strongly recommend that you create a rescue disk in case there are any problems with corrupt files in the operating system. It is a good practice to ensure that you have this disk available.

NOTE

A rescue disk is unique to the computer for which it was created. Therefore, a new one must be made for each computer in service.

Software Utilities

There are many good-quality utility programs available today. These programs allow the experienced user to find and correct a multitude of problems. However, caution should be used when "correcting" a problem that has been identified by the software. The software might consider something a problem simply because it does not recognize it. ("If I don't know what it is, it must be bad.") In some cases, the cure is worse than the disease. Also, keep in mind that one utility will not solve every problem. As a computer professional, you will do far better to master one good software system than to have a box full of utilities that you don't know how to run effectively. Don't forget good old MS-DOS; it is full of commands that are usually forgotten or never used.

CAUTION

Older versions of utility programs are designed to work with MS-DOS and Windows 3.x. They can wreak havoc on a Windows 95 or Windows 98 system. You must also be especially careful if you're running later versions of Windows 95 or Windows 98 that use the FAT32 file system, because most utilities are designed to handle the traditional FAT16.

You should never run any application to "tune" a system that is not specifically designed for that version of the operating system. That applies triple to advanced 32-bit operating

systems such as Windows 98 and Windows 2000. Among the handiest utilities have around are virus-checking programs that are compatible with each operating system you work with, such as disk and video display diagnostic programs.

Disassembly

Disassembling a computer is a straightforward task. In most cases, you will need to remove little more than the outer cover or shroud of the case to gain access to the memory, expansions slots/cards, and the CPU. Because there are many manufacturers, each seeking to establish its own unique marketing identity, each brand has some custom components or layout. The best strategy for efficient disassembly is to locate and use the manual that came with the computer.

Often, manuals don't provide a lot of technical information, but they usually tell you how to remove the cover. The extent to which you will have to disassemble a computer depends on the specific problem or repair. Following the following procedure will help you establish a routine for completely and efficiently disassembling most computers:

1. Make a complete backup of necessary operating-system and working files.
2. Document the system (hardware and software).
3. Create a clean work area with plenty of room and light.
4. Gather all the necessary tools for the job.
5. Implement safety procedures.
6. Turn off the computer.
7. Disconnect the power cables.
8. Wear an antistatic wrist strap.
9. Locate the screws for the cover—check the manual to discover the location of the screws (sides or back).
10. Remove the screws. It's a good idea to store them in a box or plastic tube to keep them from getting lost.
11. Remove the cover from the computer.
12. Document the location of expansion cards and drives.
13. Remove all the cards and place them in antistatic bags.
14. Document the location and connections for each drive (pay special attention to the red wire on the data cables—this identifies the location of pin 1 on the device and driver).
15. Remove the data and power supply cables.
16. Remove the drives from their appropriate bays—look on their sides for the screws (check the manuals).
17. Remove the motherboard.

Reassembly

To reassemble a computer, you simply follow the same procedures as for disassembly, but in the reverse order. When installing components, remember the following:

- Do not force connectors into place—if they don't fit easily, they are probably in the wrong place.
- Expansion cards often require some force or side-to-side movement to fit into place, but do not force them.
- When removing cables, remember the pin 1 locations. Check notations on the circuit boards, and look for the red wire on the ribbon cables.
- Connect the cables to the drives *before* installing them in the bays.
- Test the system before replacing the cover.

Upgrading a Computer

Most frequently performed task by a computer professional is to upgrade old systems to the latest technologies. This ability to expand and upgrade a computer can prolong the life and utility of a system. However, sometimes even the simple addition of a new piece of software can lead to hardware conflicts and the subsequent need for an upgrade, as a computer owner tries to squeeze one more year out of "old faithful." This lesson discusses many aspects of computer hardware upgrades.

Memory, Memory, Memory

Does this computer have enough memory? This is the question that most frequently causes users to seek a computer upgrade. As programs and hardware get faster and are required to process more graphics and animation, the need for memory is as important as the need for speed. Memory upgrades are perhaps the simplest to perform, but they can be very confusing without advance planning. Purchasing the right memory for the job is more than half the process of the upgrade. Before installing memory, there are five things to consider:

- Memory chip format , Memory speed
- EDO RAM , Parity
- Cache memory

The best source of information—which should be checked *before* obtaining memory—is the documentation that comes with the computer's motherboard. This source will generally list the type of memory required, how many SIMMs are required, and their location on the motherboard. If this information is not available, open the case and look. Some documentation provides a chart that includes exactly what memory has been installed and what is needed to upgrade to a given level. The following table gives you an idea of the kind of chart you might come across.

On-board	Bank 0	Bank 1	Total
8 MB			8 MB
8 MB	4 MB	4 MB	16 MB
8 MB	8 MB	8 MB	24 MB
8 MB	16 MB	16 MB	40 MB
8 MB	32 MB	32 MB	72 MB

Disabled	64 MB	64 MB	128 MB

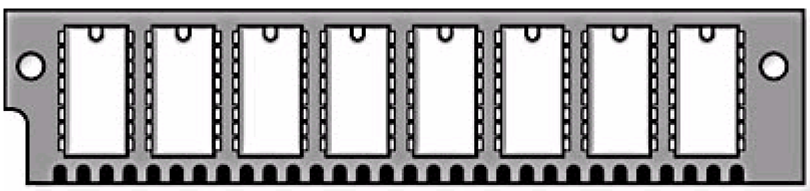
NOTE

You can add memory with SIMMs (single inline memory modules) or DIMMs (dual inline memory modules).

SIMM Formats

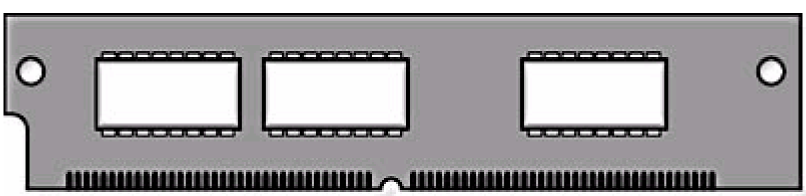
SIMMs are provided in two basic, physical formats: a 30-pin and a 72-pin chip. Format is the first consideration, because the chips must fit into the motherboard. This configuration, along with the size of the processor, determines how many SIMMs are required to fill one bank.

The 30-pin formats contain memory in 8-bit chunks. This means that a 32-bit processor requires four SIMMs to fill one bank. Typical 32-bit processors consist of two banks of SIMMs and therefore eight slots.



30-pin SIMM

A 72-pin format is larger and supplies memory in 32-bit chunks. Only one SIMM is required for a 32-bit machine. A Pentium processor has a 64-bit data path and requires a 72-pin SIMM.



72-pin SIMM

Memory is normally sold in multiples of 8 MB. However, some machines will have 8 MB of "on-board" memory (usually soldered in place on the motherboard). When memory is soldered in place, it cannot be changed but this should not be considered a problem; it can be disabled. A computer equipped with this on-board memory can provide 8 MB of memory to the system without having any SIMMs installed in the slots.

For such computers, installing 16 MB of RAM too would yield a total of 24 MB of RAM; if 64 MB were to be added, the total RAM would be 72 MB, and so on. In general, the idea of hardwiring memory on a system has died out on desktop PCs.

DIMM Formats

DIMMs are much easier than SIMMS to install or remove, because they require only one card, which is simply pushed into a module slot. The "key" cut into the edge that goes into the slot prevents the card from being inserted the wrong way. The one problem you face is choosing from the wide variety of memory types available. When ordering a new DIMM, you must know exactly the memory type supported by the system. DIMMs are found in larger memory sizes than SIMMs, ranging to 256MB and beyond for single cards.

Memory Speed

Memory speed is the amount of time required to access data and is measured in nanoseconds (ns); each nanosecond equals one billionth of a second. Two important considerations arise when addressing memory speed:

- The lower the number, the faster the chip speed
- All chips in the same computer should run at the same speed.

Typical chip speeds are 50, 60, 70, and 80 nanoseconds. Be sure to check the motherboard documentation or the existing chips to determine the correct speed to use.

EDO RAM

The EDO RAM (extended data out read only memory) chip is used extensively with Pentium processors. This chip can improve read times and overall performance by up to 30 percent. This performance gain is possible because the chip continues to output data from one address while setting up a new address.

Cache

Cache memory can be found as either L1 or L2. The L1 cache is built into the processor and cannot be changed. The L2 cache, on the other hand, can be either built into the processor, or built onto the motherboard, sometimes both. In most cases, cache memory is fixed, but some machines allow L2 cache to be upgraded or expanded. Cache memory is sometimes found on older motherboard (as DIPPs—dual in-line packages). Check the motherboard documentation to determine what, if any, upgrades can be made to the cache.

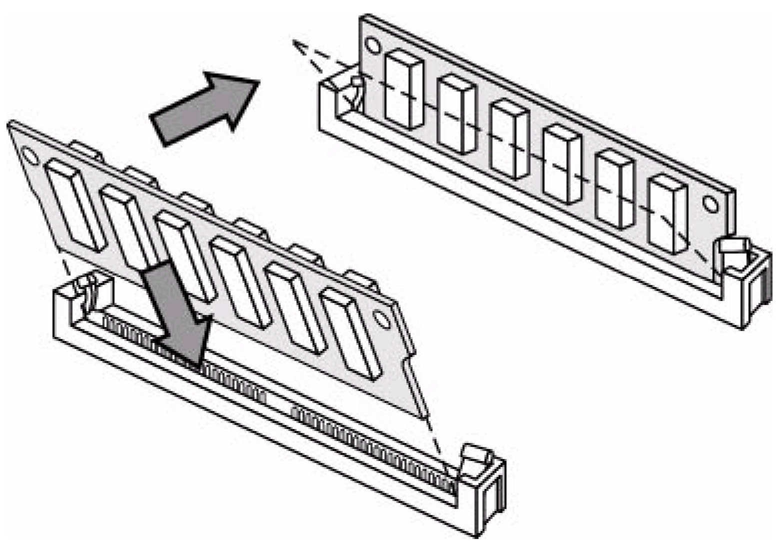
IMPORTANT

You need to take special care when installing DIPP chips. They are sensitive to ESD, can easily be installed backwards (look for pin 1 alignment), and the pins can be broken or bent during insertion.

Installing RAM

Installing RAM is a simple process. The only problem is that the slots are not always easily accessible. Sometimes you will need to relocate wires temporarily or even remove expansion cards. This simple procedure usually works:

1. Turn off the computer.
2. Disconnect all external devices (AC power and monitor power).
3. Follow the appropriate ESD safety procedures.
4. Remove the cover of the computer.
5. Locate the SIMM banks and determine that you have the correct size, speed, and quantity of SIMMs.
6. Insert the SIMMs in the slot at a 45 degree angle (backwards) and then snap it into the upright position, as shown in Figure 14.4. Be sure that the "notch" in the SIMM matches the slot. If it doesn't fit easily, it is probably installed incorrectly!



Installing a SIMM

7. When the SIMM is in an upright position, be sure that the metal retaining clip snaps into position. This clip holds the SIMM in place and must be opened before any SIMM can be removed.
8. Replace any temporarily removed or relocated wires or expansion cards. Check others to make sure they have not been loosened or disconnected.

9. Replace the cover of the computer.
10. Reconnect the power, monitor, and any other needed external devices, and start the computer.

The computer should recognize the new memory and either make the correction or automatically go to the setup program. In many cases, you need only exit setup to save the changes.

CPU Upgrades

Installing a new CPU is becoming less common as the prices of new motherboard/ CPU combinations (and even new machines) continue to drop. In many cases, installing additional memory is a more effective upgrade than installing a new CPU. Still, as a technician, you need to know how to update the CPU in an existing machine.

In many cases, upgrading a CPU is as simple as removing the old one and inserting the new one. First, you need to determine whether the CPU can be upgraded and, if so, to what? The answer to this question lies in the motherboard. The motherboard must have the appropriate socket, data bus, address bus, and crystal to support the new CPU. Consult the documentation that comes with the motherboard—this documentation usually contains a table that defines which CPUs can be installed. If you are unable to find the documentation, or the processor that you want to install is not listed (because it's of newer vintage than the documentation), you will need to consult the motherboard manufacturer either through the company's Web site or with a phone call to the company's tech support department. Be sure to check on any required jumper settings and BIOS upgrades at the same time.

A small upgrade, going from one level of the same CPU family to another, is usually no problem. But if you want to upgrade a 386 to a Pentium, or a Pentium to a Pentium III, a new motherboard is the only answer. The same is true if the CPUs are coming from different chip makers.

General Procedure for Installing a CPU

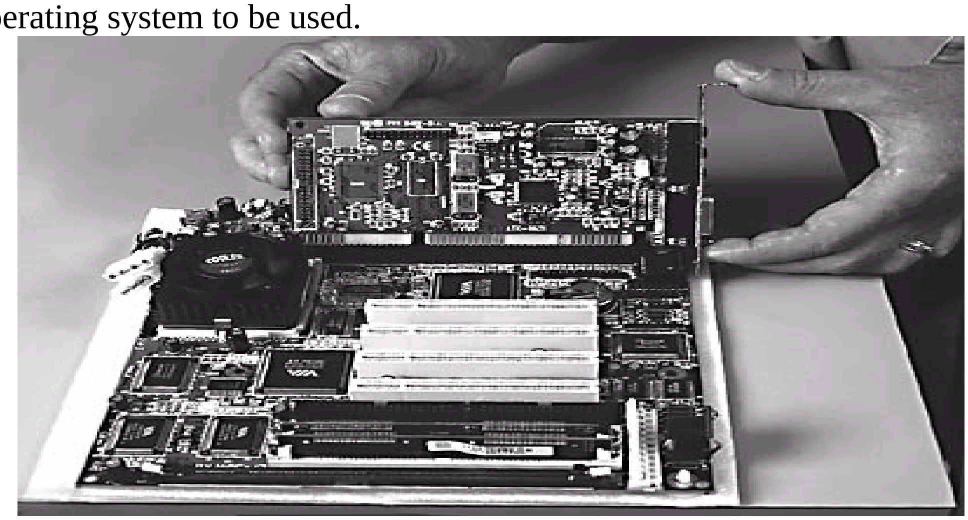
Perhaps the most difficult part of upgrading a CPU is determining the limits imposed by the motherboard. But after the decision is made and you have the new processor in hand, the actual installation is quite easy. Follow this general procedure to install a CPU:

1. Turn off the computer and unplug the power cord.
2. Disconnect external devices (AC power and monitor power).
3. Follow the appropriate ESD safety procedures.
4. Remove the cover of the computer.
5. Locate the socket for the CPU. It might be on the motherboard or on a removable processor card.

6. Remove the old processor. This may require special tools for older processors. Pentium II and III packages are Slot 1 designs, which slide into a slot much like those used for an expansion card. The original Pentiums (60-166 MHz) and Pentium Pro models usually have a ZIF (zero-insertion-force) socket. The ZIF socket is opened by moving the handle to the upright position. (This should not require force.) The CPU can then be easily removed.
7. Install the new processor. Be certain to align the chip properly (this is critical!). Pin 1 on the CPU must fit pin 1 in the socket. There are several ways to identify this pin. Various chip manufacturers, and different versions of a manufacturer's chips, use different methods to mark installation orientation. Slot 1 CPU packages, for example, have a key in the slot, and it fits only one way. Other CPUs have similar schemes appropriate to their socket design. Look for a key pinhole in one corner, a blunt edge on one corner of the socket, a dot, a corner with a pin arrangement that differs from the others, or some other identifying mark. Align this mark with the corner of the socket that contains a blunt edge. If you encounter any resistance or you have to apply any pressure when inserting the CPU, recheck the chip's orientation and alignment, and reinsert the chip. After the chip is in place, secure the ZIF handle. You might need to check the documentation to make sure the chip is installed correctly.
8. Set any jumpers or switches on the motherboard. Check the documentation.
9. Replace the cover and power up the computer.
10. Reconnect any peripherals (keyboard, mouse, monitor).
11. Make changes to the CMOS setup, if required.

Expansion Cards

Installing an expansion card is one of the most common system upgrades. Adding faster video cards, more ports, or improving sound quality are common reasons for plugging in a new card. Before installing (or purchasing) an expansion card, it is a good idea to make sure it will work in the system to be upgraded, and that appropriate drivers are available for the operating system to be used.



Motherboard and expansion card

Ask these questions first:

- Is adding a new card the most cost-effective way to make this upgrade, given the type of device and performance/capacity desired? In some cases a USB (universal serial bus) peripheral can offer the same features without requiring the case to be opened.
- Are there any expansion slots available? If no slots of the type required are available, you will have to make some room. To do that, you will need to do one of three things: replace separate, single-port cards with one multifunction card that provides all port connections; use a SCSI card and a chain of SCSI devices, if available; or use USB. if available.
- Will the card fit in the type of slot available? Does it match the bus type of the motherboard?
- Are there any available I/Os and IRQs in the system? If so, write them down.
- Is there enough memory (RAM and hard-disk) available to run the device and its software?
- Does the card require a DMA (direct memory access) channel? If so, is one available?
- What are the potential conflicts with other cards and devices?
- Will the software operating system support this card? If so, are all the necessary drivers provided with the operating system or will they have to be provided with the card?

After you have determined that the expansion card will work, the installation is a simple three-step process:

1. Set any jumpers or switches for IRQ and I/O addresses.
2. Install the card and cables.
3. Install any software for the card.

Step 1 (IRQ and I/O setup) is perhaps the most confusing and frustrating part. This is especially true if the computer has not been properly documented.

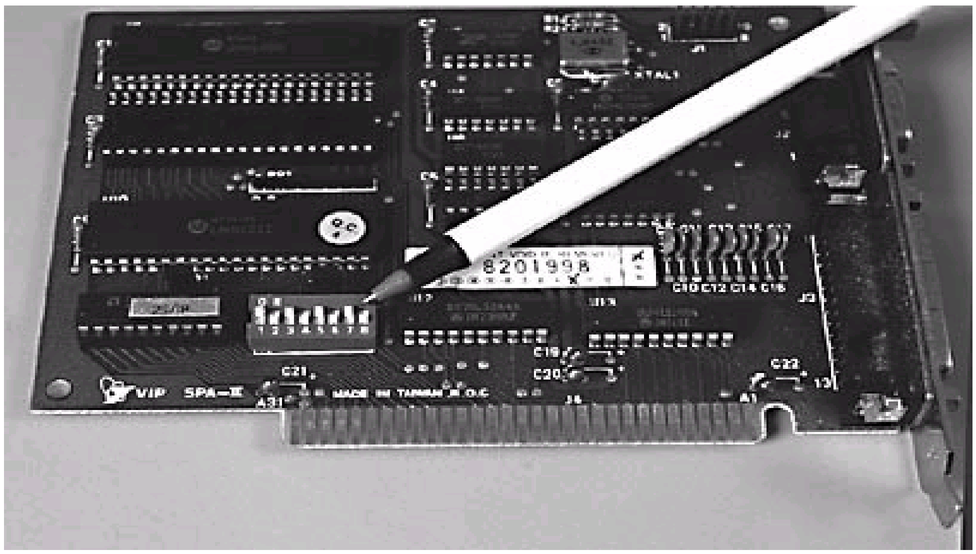
IMPORTANT

Be sure to cover all slots. A missing expansion-slot cover can cause a computer to overheat. Be sure that no screws or other items are left loose inside the case that could possibly short out a component after the case is closed.

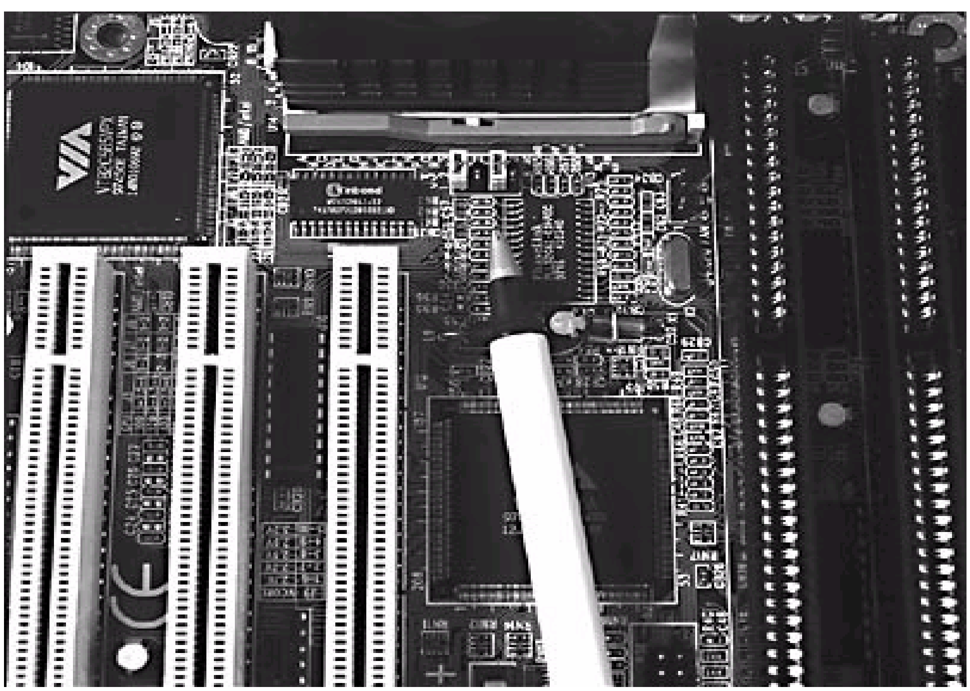
General Procedure for Installing an Expansion Card

A card that is not Plug and Play capable is a bit more complicated than one that is.

1. Read the documentation that comes with the card and note any special requirements or limitations *before* you start the installation.
2. Check the computer documentation (run MSD or other diagnostic program) and determine which IRQs and I/O addresses are available. Check the computer documentation (run MSD or other diagnostic program) and determine which IRQs and I/O addresses are available.
3. Configure any jumpers or switches on the card. (See Figures 14.6 and 14.7.) Note that some cards might require changes in order to prevent conflicts and allow all devices to work.

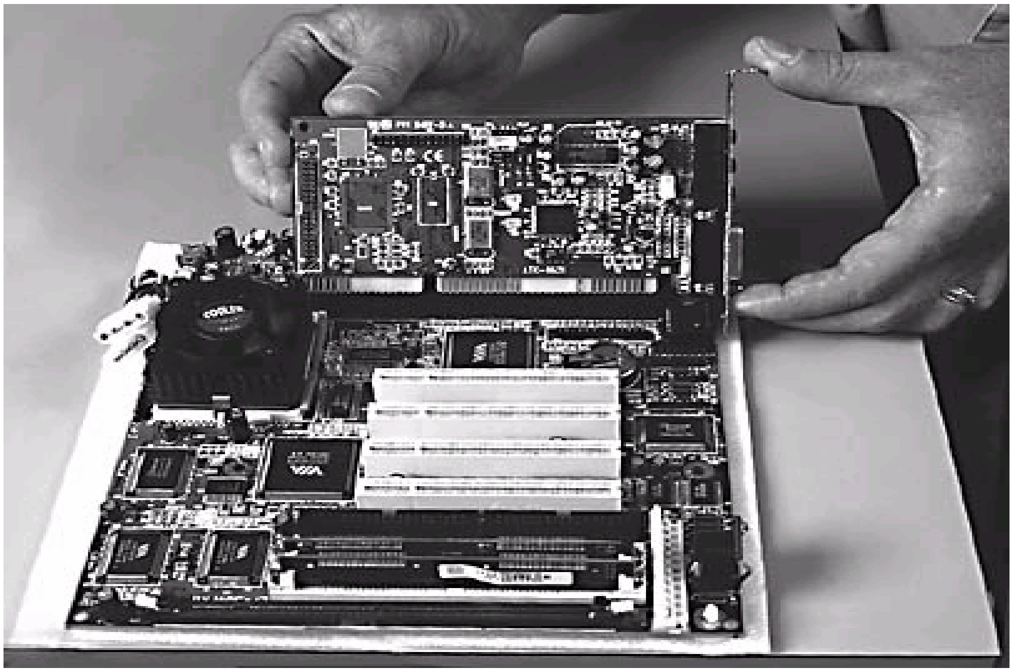


Switches



Jumpers

4. Turn off the computer and unplug the power cord.
5. Follow the appropriate ESD safety procedures.
6. Remove the cover of the computer.
7. Install the card in a free slot. Power up the computer, note any conflicts, and make adjustments as necessary. Remember to remove power and use ESD precautions when making changes.



Installing an expansion card

8. Replace the cover.
9. Install any software drivers or applications.

General Procedure For Installing An Expansion Card

The latest technology available for installing expansion cards is called Plug and Play. This is an independent set of specifications developed by a group of hardware and software companies that allow the user to make configuration changes with minimal adjustment. Simply install the card, turn on the computer, and use the device.

In order for Plug and Play to work, the device must be able to identify itself and its system requirements to the system. The operating system will then set the device and make any other adjustments, such as reconfiguring other devices, as required. In order for a Plug and Play device to work immediately, you must be sure that the computer hardware—motherboard, BIOS, and other components—the operating system, and the device are Plug and Play-compliant.

Drives

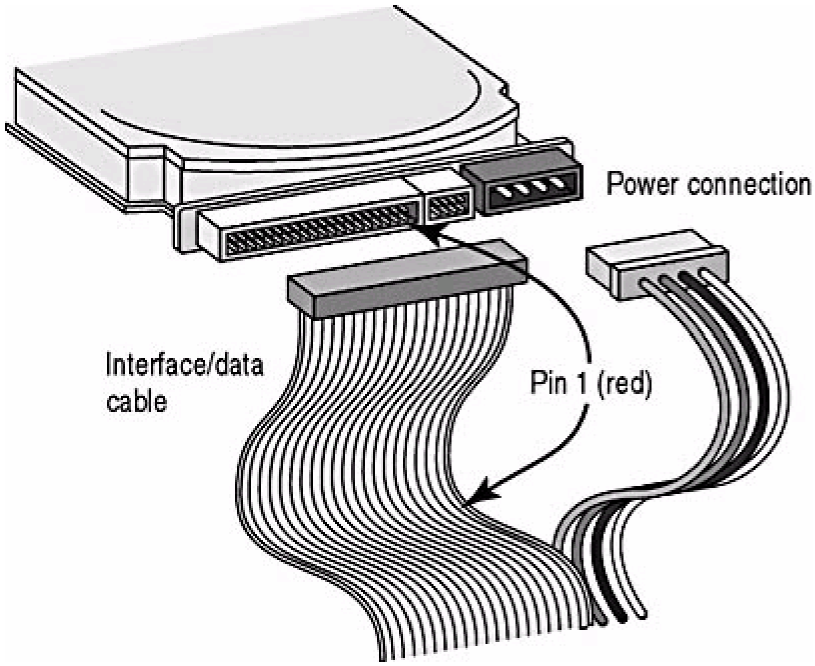
Installing a new drive is not difficult. However, you must answer a few questions before purchasing a new drive:

- Will the drive physically fit inside the computer? Some desktop cases have only enough space for one hard disk drive, or the available space might be occupied by another device (a CD-ROM drive or floppy disk drive). If there is not enough space, you will have to consider alternatives such as external SCSI (Small Computer System Interface) drives, USB drives, and parallel port interface drives.
- Will the computer's BIOS and operating system support the size (storage capacity) of the drive?
- Will the drive controller support the new drive? A second or updated controller might be required.
- Are there sufficient cables (data and power) to install the drive?

Procedure for Installing an IDE Drive

Installing an IDE (Integrated Drive Electronics) drive will require some hardware and software preparation to get it running properly. Hardware preparation includes ensuring that you have the correct drive, a place to physically install it, and the proper cables to connect it. Software preparation includes at least a bootable MS-DOS disk containing a minimum of FORMAT and FDISK. A Windows 95 or 98 startup disk will do the job. If you don't already have such a disk, be sure to create one before removing the old drive. Follow these steps to install the drive:

1. Collect all the necessary documentation for the drive and the computer.
2. Turn off the computer and unplug the power cord.
3. Follow the appropriate ESD safety procedures.
4. Open the case of the computer.
5. Set the jumper for the drive. Consult the documentation that came with the drive. It must be set to single use or master or slave. If this is a second drive, both drives might require jumper settings.
6. Connect the cable to the drives. The end connector should be plugged into the master drive. Be sure the cable orientation is correct (pin 1 goes to the red wire).
7. Connect the power cable.



Cable Connections

8. Install the drive in its bay.
9. Reconnect the power, boot up the computer, and run the CMOS setup utility. The CMOS must be set to recognize the new drive. See the manufacturer's documentation for proper configuration. Don't forget that drives larger than 528 MB (approximately) will require logical block addressing (LBA).

Some larger drives installed in older machines might require the use of disk-management software. This software is usually provided with the drive (in some cases, it's already loaded in the drive). Follow the instructions provided by the manufacturer to extract and use this software. It's especially important to document the use of such software, and to make sure it is included on any boot/rescue floppy you prepare for such a system.

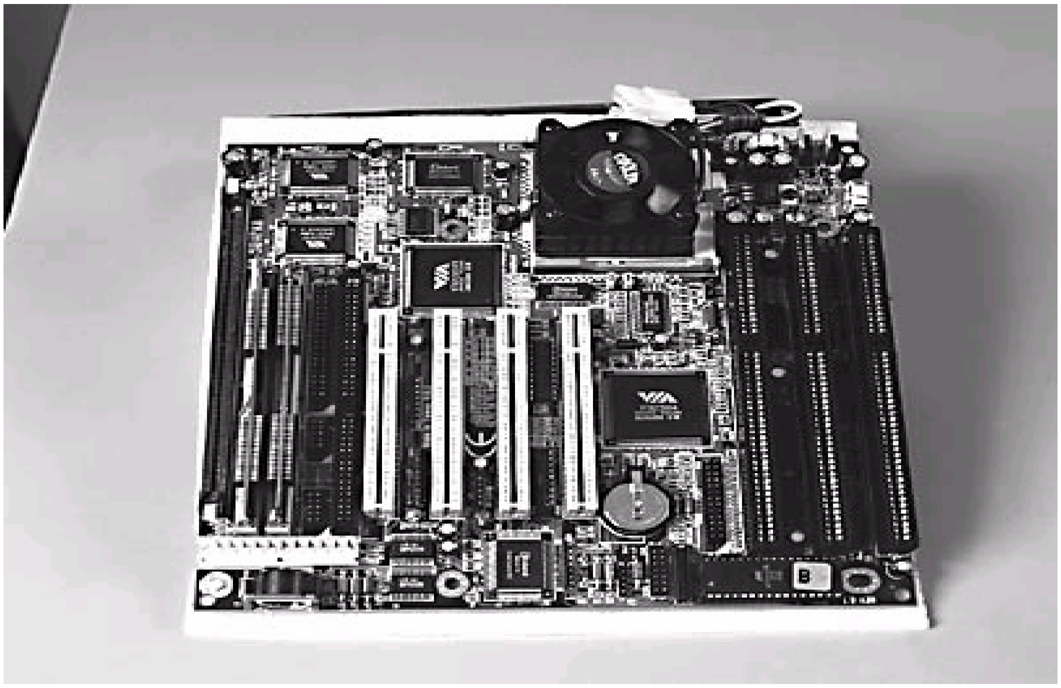
To complete the installation:

1. Boot the computer from the bootable floppy disk and run FDISK to set the partition (or partitions).
2. Format the drive. If it is the only drive or the bootable drive, it must be formatted with the system files.
3. Replace the cover of the computer.

The drive is now ready for software to be installed.

Motherboards

Installing a new motherboard is one way to completely overhaul a computer. In many cases, it is the most inexpensive way to get a new computer. This type of upgrade usually works best with IBM clones. Many of the larger manufacturers use proprietary motherboards that can only be replaced with one made by the same manufacturer to ensure compatibility with other components.



Motherboard

Before deciding to undertake this major overhaul, there are several questions to answer:

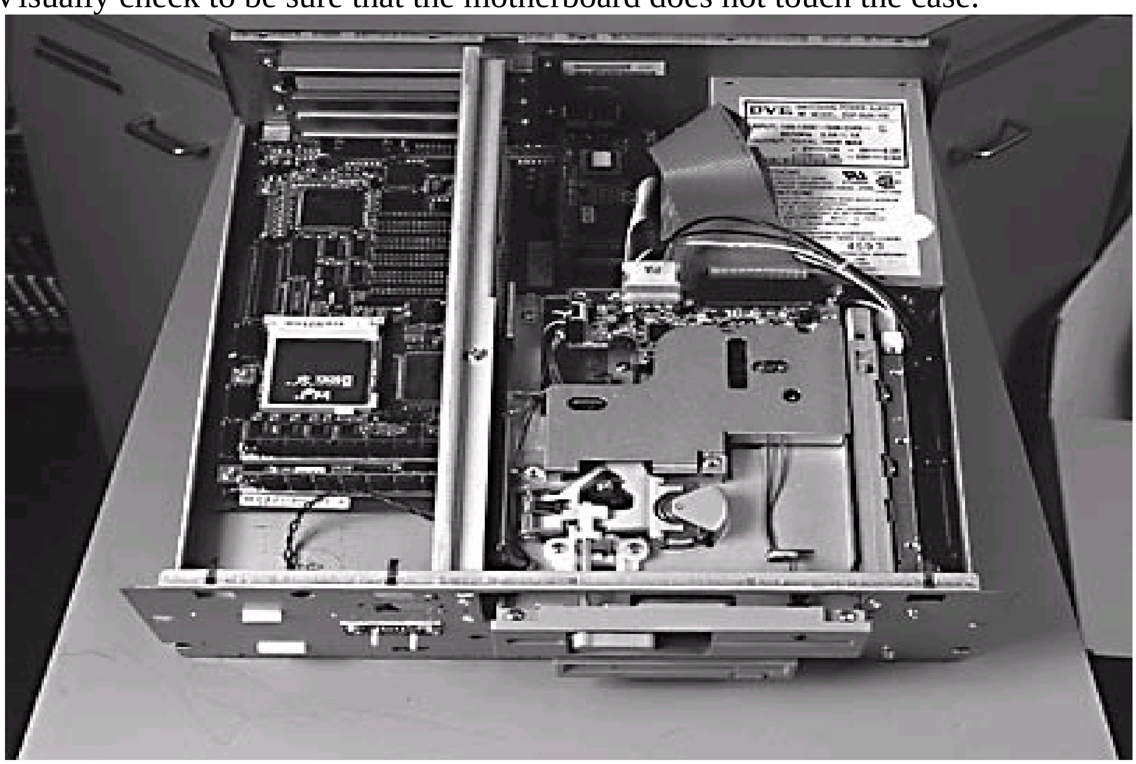
- Will the motherboard fit into the existing case? Check the size and alignment of mounting holes (plastic standoffs that keep the circuitry from contacting the case).
- Does the motherboard have the same built-in COM and LPT ports?
- Does the motherboard have a built-in video card?
- Will the existing expansion cards fit the motherboard's expansion bus slots? Are there enough slots available to accommodate the existing cards? The expansion slots should go toward the back of the computer, where the openings are located.
- Is the power connector located on the same side as the power supply? It should be as near to the power supply as possible.
- Will the existing drives (CD-ROM, IDE, or SCSI) work with the controllers on the motherboard?
- Will the SIMMs on the old motherboard work with the new motherboard?
- Will the upgrade meet your current and future requirements?

Installing a new motherboard is a major task and requires complete disassembly, reassembly, and setup of the computer and all its devices. You will put everything covered so far in this lesson into practice when you replace a motherboard. The best advice is to prepare everything ahead of time and to take good notes while disassembling the old motherboard.

Replacing a Motherboard

Replacing a motherboard is probably the most difficult task (from the perspective of physically replacing parts) that a computer technician will take on. It amounts to no less than building a computer, because it will, in many cases, require complete disassembly first. Although complex, if you carry out the following procedure one step at a time, you should not face any problems.

1. Complete an installation checklist and make sure all the necessary parts are available and will fit into the computer.
2. Follow the "disassemble a computer" steps set out in Lesson 1; these include removing all the screws and standoffs (taking care to keep them in a safe place) and the old motherboard.
3. Check all the settings on the new motherboard.
4. Install the new motherboard. Position the standoffs and make sure they align with the case and the motherboard. Verify that the motherboard is positioned correctly (expansion slots facing the back of the computer). Carefully tighten the screws. Visually check to be sure that the motherboard does not touch the case.



Motherboard in case

5. Reconnect the case switches. Use the notes taken during disassembly to verify they are in the right place.
6. Follow the steps for computer reassembly set out in Lesson 1.
7. Test to be sure that the computer boots up.
8. Complete the final testing and close the case.

Chapter 15

Software Operating Systems

OBJECTIVES

Operating System Basics

MS-DOS

Operating System Basics

MS-DOS

The term *software* refers to any program (set of instructions) that directs a computer to carry out a task or function. Software can be divided into two categories: operating systems and applications. Operating-system software is used to manage hardware, data, and application software. No computer can run without an operating system. MS-DOS, Windows 3.x, Windows 95, Windows 98, Windows NT, Windows 2000, OS/2, and UNIX are examples of operating systems.

DOS stands for disk operating system. The full proper name is preceded by the name of the manufacturer—for example, MS-DOS (for Microsoft and also used by "clone" makers) or PC-DOS (for versions specific to the IBM Personal Computer). This version was also sometimes referred to as IBM-DOS. All versions of DOS are actually the product of Microsoft development. The modern computer professional must be familiar with MS-DOS because MS-DOS remains at the heart of Windows. Although increasingly more configuration can be done through the Windows interface, there comes a time when every computer professional is faced with a screen of MS-DOS commands. Windows 98—the most current operating system—is more independent of MS-DOS than prior versions of Windows. However, every PC still requires AUTOEXEC.BAT, CONFIG.SYS, and various initialization files. So there remains a need for the computer professional to be familiar with MS-DOS.

DOS has been produced by three suppliers, and the versions are so similar that they are considered as one operating system. The three brands of DOS are:

- **MS-DOS**, produced by Microsoft—the most popular
- **PC-DOS**, produced by IBM
- **DR-DOS**, produced by Novell

Proprietary DOS

Many computer manufacturers, especially in the early days of the 8088 and 80286 processors, produced their own versions of DOS. These were actually tweaked versions that contained additional commands or utilities specific to their hardware. Such versions might be required to run these machines and will require software that is compatible as well.

DOS was originally designed to load an operating system from a floppy disk to a computer with no hard drive (before hard drives were common). Today, any reference to DOS is synonymous with MS-DOS.

The following table provides a version history of MS-DOS.

Some computer manufacturers have produced OEM (original equipment manufacturer) versions of MS-DOS. For the most part, these versions have been cosmetically modified for specific use with their computers (such as a "splash screen" that shows the company logo during boot up). Care should be taken when using any computer with one of these operating systems—some incompatibilities with software written for MS-DOS might exist. You should be careful when substituting a generic DOS for MS-DOS if you are reformatting a hard disk drive or reinstalling the operating system. Make such a substitution only after you have verified that the generic version is compatible with all of the computer's hardware and software; the generic version might not provide the necessary support.

MS-DOS includes three programs that are required to make a drive or floppy disk bootable:

- **IO.SYS:** The interface between the hardware and the operating-system code
- **MSDOS.SYS:** The main operating-system code
- **COMMAND.COM:** The interface between the user and the operating-system code

NOTE

IO.SYS and MSDOS.SYS are hidden files that are not usually visible when you examine a disk directory.

In addition to the three core files, MS-DOS uses two other startup files. These files are not required to start the machine, but they add any additional startup configuration required by the user or applications. These files are:

- **CONFIG.SYS:** Loads extra hardware and device drivers not built into the IO.SYS.
- **AUTOEXEC.BAT:** Loads terminate-and-stay-resident (TSR) programs selected by the user and sets up the environment variables such as TEMP and PATH.

NOTE

If a user complains that a machine is not loading MS-DOS, especially after installing new hardware or software, disable these two files as a first diagnostic step. In many cases, drivers or procedure calls are conflicting or corrupting system memory.

Windows 95 and Windows 98

Unlike earlier versions of Windows, Windows 95 and Windows 98 are true operating systems.

Windows NT

In the summer of 1993, Microsoft released another operating system primarily for the professional market. When it was first introduced, it was used in relatively simple network installations. Over several revisions, it has been enhanced to support the needs of corporations ranging in size from small to large and, now, the needs of the Internet and intranets. As an operating system, Windows NT provides a high degree of networkability as well as high security levels not available in other operating systems. Windows NT uses the same GUI as Windows 3.x; therefore, making it an easy upgrade for anyone proficient in using the Windows operating system.

There are several versions of Windows NT, ranging from 3.0 through 4. In addition, there are Service Packs; inline fixes that do not change the version number. Windows NT provides three levels of operating system in each of the later versions:

- **Workstation:** A powerful, robust operating system with limited networking to allow sharing of printers and files for the professional user.
- **Server:** A complete LAN host with a variety of sophisticated features for managing users and access to printers, files, RAID installations, and other shared resources.
- **Advanced Server:** The enterprise edition that adds tools for complex network environments.

Windows 2000

The most recent addition to the Windows family, Windows 2000, is the replacement for Windows NT, adding Plug and Play support, better multimedia tools, and advanced Internet support. Like Windows NT, it comes in three versions: Workstation, Server, and Advanced Server.

NOTE

Be careful when upgrading to Windows NT or Windows 2000. They are not merely more powerful versions of Windows 98, but more robust, completely new environments. Not

all applications and hardware are compatible with them. Be sure to consider all the advantages and disadvantages before making the decision to upgrade. Then check the Microsoft Web site for the latest version of the compatibility list.

MS-DOS

The Boot Process

In computer terms, *boot up* means to start or get going. Once POST is complete (the basic hardware is determined to be OK), the computer loads software that tells it how to run. This software is the operating system (OS). The boot sequence for operating systems will vary. Several steps are involved in the MS-DOS system startup process:

1. The ROM BIOS looks for an operating system. It checks for the presence of IO.SYS and MSDOS.SYS, first searching the A (floppy disk) drive and then the C drive.
2. The operating system processes the CONFIG.SYS file (if present). The CONFIG.SYS contains information to configure the system environment including special memory management overlays and hardware drivers.
3. COMMAND.COM is loaded.
4. The operating system processes the AUTOEXEC.BAT file (if present). The AUTOEXEC.BAT loads programs and user-defined settings.
5. If no programs (such as Windows) have been invoked, COMMAND.COM presents the active-drive prompt and waits for a command.

The process of starting from a no-power condition is called a *cold boot*. Occasionally, a system might require a reset—for instance, when the computer locks up or runs out of memory. Resetting can be accomplished without turning off the computer by holding down the CTRL, ALT, and DEL keys at the same time. This is called a *warm boot*.

MS-DOS Terms and Commands

Key MS-DOS Terms

As with any operating system, it is important to understand the language or terms used. As mentioned, this is not a complete course in MS-DOS and there is no substitute for an MS-DOS manual. The following table presents some of the important terms and concepts you might encounter when working with MS-DOS.

Term	Description
.bat	An extension that indicates a batch file.
.com	An extension used to identify a file as an executable program file.
.exe	An extension used to identify a file as an executable program file.
Backslash (\)	A symbol used to separate each directory level, for instance C:\Windows\Utilities. For this reason, it is a reserved character and cannot be used as part of a file name.
Case sensitivity	The ability of the operating system to distinguish between uppercase and lowercase letters. MS-DOS commands are not case-sensitive. Traditionally, MS-DOS commands have been represented in documentation as uppercase. While you can type MS-DOS commands in either upper- or lowercase, we are representing them in this book as uppercase.
Cursor	Anytime you are entering data, whether in an application or in an MS-DOS command, the cursor (usually a small flashing line) indicates the place where the next character will be inserted. It is a good idea to always know where your cursor is.
Default drive	Each drive in a computer has its own letter designation. The default drive is the active drive. Unless otherwise specified, any commands act upon the default drive. The current default drive is indicated by the MS-DOS prompt. For example, if you want to see a directory (the command is DIR) of files on the A drive and the default drive is C, you need to type `DIR A:` otherwise you will see a directory of the C drive.
Directory	Directories—known as folders in the Windows and Macintosh environments—are locations for storing files. Every disk contains a main directory known as the *root directory*. Below the root directory is a hierarchical structure of other (sub)directories.
MS-DOS prompt	The MS-DOS prompt usually displays the active drive letter (for instance, C:) and directory. This indicates that the operating system is ready to accept the next command. (The prompt is user-changeable.)
Drive	MS-DOS assigns letters to each drive during the boot process.

pointers	
Entering Commands	You can type a command and press ENTER to execute it. If you make a mistake, correct it by using the BACKSPACE or DEL keys. Use ESC to start a command again. Use the F3 key to repeat a command.
Error messages	Brief technical messages that are displayed when an error occurs.
Filenames (also filespecs)	A filename is made up of three parts—a name of up to eight characters, a period, and an extension of up to three characters. The name can include any number, character, or the following symbols: _()~'!%$&#. Spaces cannot be used in MS-DOS filenames. Example: MYFILE.DOC
Greater than (>)	This symbol is used to indicate that a command can be redirected to an output device. For example, to redirect the directory command to a printer, type: DIR > LPT1.
Path	The address to a file. The path consists of the drive name, the location of the file in the directory structure, and the filename. Example: C:\Mystuff\Myfile.doc.
Prompt	The command prompt—user interface provided by COMMAND.COM to signal to the user that the computer is ready to receive input (for example: C:\> or A:\>).
Switches	Many MS-DOS commands can be used with a switch ("/" followed by a letter) to invoke special functions. Because no comprehensive manuals are available after MS-DOS 5, when you follow a command with a space and "/?", a list of parameters and switches available for that command is displayed.
Syntax	Syntax is the arrangement and interrelationship of words in phrases and sentences. In computer jargon, it is the correct format in which to type a command. In MS-DOS, every letter, number, and space has a value. The most common problem when typing MS-DOS commands is adding or leaving out a letter or character. Simple typing mistakes are the most common cause for "Bad command or filename" errors.
Wildcards	The question mark (?) matches any character in a specified position, and the asterisk (*) matches any number of characters up to the end

	of the filename or extension. For example, to search for files beginning with the letter "A," the command would be DIR A*.* or A?????.* (the second command would find a file that starts with the letter A and any other five characters).

MS-DOS Commands

.Command	Function
CHDIR or CD	Changes the directory. (For example, cd\word would take you to the "word" subdirectory.)
CHKDSK	Examines the FAT (file allocation table) and directory structure on a drive, checking for errors and inconsistencies that can keep you from accessing a file. It also locates lost clusters and can convert them into files for later deletion. It can also reclaim wasted space.
CLS	Clears the screen.
COPY	Copies files or disks. To copy all files from the "myfiles" subdirectory to the A (floppy) drive, the command would be: `copy c:\myfiles\*.* a:`
DATE	Changes the system date.
DEL	Deletes files. (Example: `c:\del myfile.txt)`.
DIR	Lists a directory of files.
DIR /P	Views directories, one page at a time. (Directories can be quite long.)
DIR /W	Displays wide format in columns—only the filename is listed; not size, date, or time.
DIR /W /P	Displays large directories in columns, one page at a time.
DISKCOMP	Compares two disks. The syntax is: `a:\ diskcomp a: b: or diskcomp a: a:` (the computer will prompt you to insert the second disk to be compared).
MKDIR or MD	Makes a directory.

PROMPT	Changes the appearance of the cursor.
REN	Renames a file.
RMDIR or RD	Deletes a directory. This works only if the directory is empty of all files including hidden ones.
TIME	Changes the system time.
TYPE	Displays (types) a text file.
VER	Displays the version of MS-DOS in use.

The following table lists examples of commonly used external MS-DOS commands.

Command	Function
DISKCOPY	Makes a copy of a complete disk. Requires that both the source and the destination disk have the same format.
EDIT	This command invokes the text editor program. This program is useful for making changes to text files such as editing CONFIG.SYS and AUTOEXEC.BAT.
FORMAT	Prepares a disk for receiving files. Places a root directory on the disk.
FORMAT /S	Formats a disk as a system disk.
UNDELETE	Will (sometimes) recover a deleted file. Works only if the disk has not been modified since the file was deleted.
XCOPY	Copies the contents of one disk to another disk. Does not require both disks to have the same format. (Note that it will not copy hidden files unless you use the /h switch.)

Optimizing MS-DOS

Understanding how to configure and optimize an MS-DOS computer might seem like a waste of time in view of the newer operating systems available like Windows 95, Windows 98, Windows NT, and Windows 2000; however, one must remember that MS-DOS is the basic building block of the operating environment for the vast majority of

IBM-compatible computers. All IBM-compatible computers today have evolved from this standard. The limitations imposed by MS-DOS and the need to provide backward compatibility have a lot to do with how today's computers operate.

Configuring MS-DOS

The job of configuring how MS-DOS runs on a specific computer is done by modifying the CONFIG.SYS and AUTOEXEC.BAT files. During the boot process, MS-DOS looks for these files and executes any commands or loads any drivers properly called for within them. The CONFIG.SYS file is run first and sets up and configures the computer's hardware components. The AUTOEXEC.BAT file executes commands and loads TSR programs.

CONFIG.SYS Commands

The following table lists several CONFIG.SYS settings and their functions.

Setting	Function
BUFFERS	Allocates reserved memory for transferring information to and from the hard disk.
COUNTRY	Enables MS-DOS to use country conventions for times, dates, and currency. Example: COUNTRY=044,437,C:\DOS\COUNTRY.SYS (You do not need to use this in the United States unless you wish to use an alternate convention.)
DEVICE	Loads a device driver into memory. `DEVICE=C:\DOS\EMM386.EXE`
DEVICEHIGH	Loads a device driver into upper memory.
MS-DOS	Loads part of MS-DOS into upper memory area. `DOS=HIGH, UMB`
FCBS	Specifies the number of file control blocks (FCBS) that MS-DOS can have open at the same time.
FILES	Specifies the number of files that MS-DOS can hold open concurrently. `FILES=60`
INSTALL	Loads a memory-resident program. `INSTALL=C:\DOS\SHARE.EXE/F:500 /L:500`

LASTDRIVE	Specifies the maximum number of drives the computer can access. `LASTDRIVE=Z`
MOUSE.SYS	Loads a mouse driver.
NUMLOCK	Specifies whether the NUM LOCK key is on or off when MS-DOS starts.
SHELL	Specifies the name and location of the command interpreter. The interpreter converts the typed command to an action. The default for MS-DOS is COMMAND.COM.
SWITCHES	Specifies special options in MS-DOS. The /N switch will disable the use of the F5 and F8 keys to bypass startup commands (used for security).

Here is a sample CONFIG.SYS listing:

```
DEVICE=C:\WINDOWS\HIMEM.SYS
DEVICE=C:\DOS\EMM386.EXE NOEMS
DOS=HIGH, UMB
BUFFERS=20
FILES=80
FCBS=4,0
LASTDRIVE=Z
DEVICE=C:\CDROM\NEC2.SYS /D:MSCD001 /V
```

AUTOEXEC.BAT Commands

The following table lists several commands that you'll find in the AUTOEXEC.BAT file.

ECHO	Displays commands as they are executed. @ECHO OFF: Suppresses the display of commands as they are executed.
PAUSE	Stops the execution of AUTOEXEC.BAT and displays the message "Strike any key to continue."
PATH	Defines the search path for program commands.
SET	Displays, sets, or removes MS-DOS environment variables.
SMARTDRV	Provides disk caching.
KEYB	Configures a keyboard for a specific language.

SHARE	Starts the Share program, which will install the file sharing and locking capabilities.
DOSKEY	Loads the DOSKEY program. You can use the DOSKEY program to view, edit, and carry out MS-DOS commands that you have used previously.
MOUSE.EXE	Loads a mouse driver.
PROMPT	Sets the display of the command prompt.

Chapter 16

Windows 98 and Beyond

OBJECTIVES

GUI Operating System

Installing & Configuring Windows 98

How Windows Works

GUI Operating System

When you first boot up Windows 98, you see the message "Starting Windows 98." At this point, Windows is starting the DPMI. After the DPMI is loaded, Windows 98 loads the GUI. Notice that it is not necessary to use the GUI to boot up to Windows 98; this is important because many computer repair functions, particularly for the hard disk drive, are handled at a MS-DOS prompt.

The Windows 98 GUI

The Windows 98 GUI is a protected-mode. It loads its *own* device drivers. It does not need to load device drivers from CONFIG.SYS unless Windows 98 does not have a built-in protected-mode driver for a particular device. The GUI loads a protected-mode driver for most CD-ROMs, has VCACHE for disk caching, and protected-mode mouse support for Windows 98, Win 3.x, and MS-DOS applications. As with CONFIG.SYS, there is no requirement for AUTOEXEC.BAT.

The Virtual Memory Manager (VMM) supports memory usage at the GUI. When the GUI is loaded, VMM takes advantage of the power of 386 (and better) protected mode to create virtual machines (VMs)—one for Windows 98 and one for any MS-DOS program running in Windows 98. Along with the VMM is the Installable File System (IFS), which provides support for hard drives, CD-ROMs, and network drives. The IFS also provides the support for long filenames. The IFS runs both for MS-DOS 7 and the GUI.

While the GUI is running, the main functions of Windows 98 are handled by three core operating components:

- **The kernel:** loads applications, handles all aspects of I/O, and allocates virtual memory.
- **The user component:** handles all input from the user—mouse, keyboard and so forth.
- **The graphical device interface (GDI) modules:** provide routines that draw the images on the display.

Windows 95 applications use the operating system as a pool of resources (memory, modem, video, and printer). Programs or dynamic-link libraries (DLLs) make "calls" to these three modules whenever they need to place something on the screen, check the status of the mouse, use memory, or gain access to anything else they might need. This operates similarly to Windows 3.x. However, there are also major improvements:

- Most functions run in full 32-bit protected mode. Win 3.x functions run in 16-bit real mode.
- Windows 95 runs as a preemptive multitasker, allocating time to each program. Win 3.x offered a cooperative multitasking feature that made it possible for programs to take turns using processor time.
- In Windows 95, GDI and the user component use fixed memory areas called heaps to keep track of all resources. Heaps for Windows 3.x are 64 KB in size. Improvements with hardware and applications have proven that 64 KB is often too little and frequently causes problems. Windows 95 heaps are variable in size and no longer subject to this restriction.
- The user interface is at the top of the Windows 95 architecture. Windows 95 can use either its default interface, the Windows 3.x interface, or another interface. Old Windows and MS-DOS applications run within the GUI.

The Windows 98 File Structure

Disk structures have not changed in Windows 98. Partitions are still exactly the same, so the Windows 95. Windows 98 uses 32-bit FAT (FAT32) formats for hard disk drives and floppy disk. The FDISK and FORMAT commands are still used to set up hard drives. A hard drive that was partitioned with FDISK and formatted with an older version of MS-DOS is exactly like a Windows 95 drive. Windows 95 OSR2 and Windows 98 support 32-bit FAT (FAT32). This file system improves storage efficiency by reducing cluster sizes on large partitions. The following table shows the Windows 98 versions and their appropriate file system.

Version	Description
Version A	The original Windows 98 release—intended for upgrading from Windows 95. Uses FAT16 file system.
Version B	The OEM (original equipment manufacturer) version—intended for installation on new computers only. Uses FAT32.
Win 98 SE	The final Second Edition version of Windows 98 (sometimes referred to as Version C)— intended for installation on new computers only. Uses FAT32 file system. End users cannot buy this version as a retail product.

32-Bit VFAT

Disk access is provided through the 32-bit VFAT (virtual file allocation table). Unlike the 16-bit FAT used in previous versions, VFAT is a virtual device driver that operates in protected mode. This provides more reliability and works with a greater variety of

hardware. Don't confuse 32-bit VFAT with 32-bit FAT, which has to do with how data is stored on a hard disk partition (the cluster size) while VFAT has to do with how files are accessed.

NOTE

VFAT was formerly known as 32-bit file access in Windows 3.1.

Long Filenames

In MS-DOS and Windows 3.x, filenames were in what is called an 8.3 ("eight dot three") format. That is, the filename itself was restricted to a maximum of eight characters in length, and the extension to a maximum of three characters in length. Filename and extension are connected by a period, or "dot." Windows 95 supports long filenames (LFNs). The LFN removes the 8.3 filename limitation of older MS-DOS and Windows operation systems. In a regular MS-DOS 8.3 file specification directory, all file records are stored in 32-byte records. Ten of these 32 bytes are "reserved." The other 22 bytes are used to store information on starting clusters, creation date, and creation time, and 11 bytes are for the filename itself. LFNs exist on FAT partitions by chopping the filename into 12-byte chunks (stealing one of the "reserved" bytes) and allowing up to 13 chunks, creating a filename of up to 255 characters.

When an LFN is saved, the system creates a short name that conforms to the 8.3 standard. Then, each 12 characters is cut off and stored in its own directory section. The directory entries that make up the long filename are called LFN entries. These must be backwardly compatible with MS-DOS programs and with MS-DOS itself.

NOTE

Windows takes the first six characters (no spaces) of the filename, adds a tilde (~), a number, and then the extension. If two or more files have the same first six characters, the number is incremented by one for each. For example, two files named "long file name one.txt" and "long file name two.txt" would become "longfi~1.txt" and "longfi~2.txt".

To make LFNs compatible with MS-DOS means to make sure that MS-DOS ignores the LFN entries in the directory structure. This is achieved by giving LFN entries the bizarre attribute-combination of hidden, read-only, system, and volume label. There is nothing in the MS-DOS code that tells it what to do if it runs into a file with this combination of attributes, so MS-DOS will not interfere with them.

CAUTION

Older disk utilities are incompatible with long filenames and will try to erase the LFN entries. It is critical that any disk utility that tries to diagnose the directory structure, including the SCANDISK that is included with MS-DOS 6 and earlier, should never be run on a computer with LFNs. The SCANDISK that comes with Win 95 is compatible with LFNs.

The Registry

The Registry consists of two binary files called SYSTEM.DAT and USER.DAT that are located in the Windows directory. Because they are hidden read-only files, they are prevented from being accidentally removed or changed. The Registry holds information on all the hardware in the computer, network information, user preferences, and file types, as well as anything else that pertains to Windows 95. The idea of the Registry is to have one common database for everything that comprises the computer. The Registry is designed to replace CONFIG.SYS and AUTOEXEC.BAT, as well as every .INI file (especially WINDOWS.INI and SYSTEM.INI). Windows 95 still reads these files at boot up to provide backward compatibility with Win 3.x programs that use them.

Windows 98 Enhancements

Windows 98 is dramatically different from any of the Windows series—not only in its look, but because it provides many enhancements that make it even more versatile and user-friendly. For a hardware technician, proficiency in using the Windows 98 environment requires a great deal of study and hands-on experience.

Windows 98 represents an evolutionary step in operating the system. The following table compares the features of Windows 3.x, Windows 95, and Windows 98.

Windows 3.x	Windows 95	Windows 98
A 16-bit program.	A 32-bit program.	A 32-bit program.
An MS-DOS-based application that supplements an operating system.	An operating system.	An operating system.
All hardware upgrades require a software driver to be installed in order for the new hardware to	Supports Plug and Play hardware installation. Allows user to simply plug in new hardware and let Windows 95	Supports Plug and Play hardware installation. Allows user to simply plug in new hardware and let Windows 98 configure it. Expanded hardware

work with Windows.	configure it.	support for USB (universal serial bus), IEEE1394 serial bus, ACPI (Advanced Configuration and Power Interface) and DVD (digital video disc), Windows Driver Model (WDM), Multiple Display Support, and more.
Supports filenames up to only 8.3 format, such as filename.doc or spreadsh.xls.	Supports filenames up to 250 characters.	Supports filenames up to 250 characters.
Supports only multitasking.	Supports both multitasking and multithreading (running several processes in rapid sequence within a single program).	Supports both multitasking and multithreading.
Cannot run the new 32-bit applications.	Can run older Windows and MS-DOS applications, plus the new 32-bit applications written specifically for Windows 95 and Windows NT.	Can run older Windows and MS-DOS applications, plus the new 32-bit applications written specifically for Windows 95 and Windows NT.
Program Manager and program group-centered interface.	Document and work-centered interface.	Document and work-centered interface. Optional Web-like interface.
Designed to be a stand-alone, single-user computer interface.	Designed to be an interactive terminal on a local area network (LAN), wide area network (WAN), or remote or dial-up connection.	Designed to be an interactive terminal on a local area network (LAN), wide area network (WAN), or remote or dial-up connection.
Uses Program Manager as the	Designed around a Start button used to launch a	Designed around a Start button used to launch a program or

starting point.	program or open a document.	open a document.
Windows for Workgroups provided limited support communication and networking.	Integrated support for communication and networking. Provides TCP/IP (Transmission Control Protocol/Internet Protocol) and PPP (Point-to-Point Protocol) support.	Improved support for communication and networking. Network-card friendly (supports Network Driver Interface Specification—NDIS). Supports PPTP (Point to Point Tunneling) and VPN (Virtual Private Network). Supports Windows Sockets 2.0, Web-based Enterprise management (WBEM), and much more.

Advantages of Windows 98

There are several changes, not readily visible to the user, that make Windows 98 superior to the older versions of Windows.

Windows 98 has a better system of managing computer resources: memory (RAM), video RAM, hard disk drives, and network communications.

Programs designed to run on Windows 98 carry out tasks effortlessly. They tend to be faster, more efficient, and less likely to crash than with earlier Windows programs.

Microsoft's goal was to build a backward-compatible 32-bit operating system that allowed people to continue using their favorite MS-DOS programs and existing hardware. Some additional advantages of Windows 98 are:

- **User-definable interface:** The desktop is completely and easily customized by the user.
- **New accessories:** The standard Windows accessory package has been improved and expanded.
- **Interconnectivity:** It is designed to run as part of any network system.
- **Direct cabling:** Any two computers with Windows 98 and a free COM or parallel port can be connected for data sharing. One computer acts as a host and shares its files with the other.

Installing and Configuring Windows 98

Installing Windows 98

Windows 98 can be installed as upgrade to Windows 95; Clean installation on a new machine; or as a dual-boot system with either Windows XP, Windows NT, or other operating systems like Linux. Each method has unique requirements and limitations. The basic process is the same for all installations, but you must be sure that you understand the differences that do exist so that the final result is the solution you or your client needs. In many cases, you might have to compromise to achieve the best results.

Hardware Requirements

It is strongly recommended that a Windows 98 installation be performed from CD, not from over 70 floppy diskettes. Also, Windows 98 performance will be very poor unless it is installed on at least a Pentium 200Mhz machine with at least 32Mb RAM. Windows 98 does consume more resources than Windows 95 and so it functions better on a machine with more power. Windows 98 is generally better supported by software vendors than Windows 95. Lastly, it is always better to use Win9x drivers for sound cards, network cards, modems, video cards, etc. than using DOS 16-bit ?real-mode? drivers.

- Preparing the Computer for Installations? Plan whether a fresh install or an upgrade is going to take place.

- The Windows Startup Disk

For a fresh install, the hard disk will need to be partitioned using FDISK.EXE and then formatted using FORMAT.EXE. In order to execute these commands, if the hard drive has no operating system on it, one should boot the machine using a DOS bootable diskette, or a Windows 9x Startup Diskette. This Startup Diskette comes with a new purchase of Win9x, or one can easily create one. To create a Windows Startup Diskette, you will need to use a machine that already has Win9x installed on it. Then go to Start, Settings, Control Panel and double-click the Add/Remove Programs icons. Select the Startup Disk tab and click on Create Disk. Once this diskette is created, you must then ensure that the startup disk is capable of accessing the CD-ROM installed on your machine. The Windows 9x Startup Disk automatically loads standard CD-ROM drivers, but if this does not work for your particular CD-ROM, then you can edit the Autoexec.bat and Config.sys files to reflect the type and manufacturer of the CD-Rom you do have. For example:

CONFIG.SYS:

Files=25

Buffers=9,256

DEVICE=C:\PANCD.SYS /B:25 /N:PANCD001

- OR-

DEVICE=C:\OAKCDROM.SYS /D:CDTHING

AUTOEXEC.BAT

PATH=C:\;C:\DOS

MSCDEX.EXE /D:PANCD001 /L?D M:100

- OR -

MSCDEX.EXE /D:CDTHING

After the basic Windows 9x Startup Disk is created, you may also want to add more tools, utilities and programs to the diskette, such as XCOPY.EXE, DELTREE.EXE, etc.

Partitioning Using a Windows 98 Boot Disk

With the Windows Startup Disk in a floppy drive, once can now boot the computer to an A:\prompt. Now the hard drive can be partitioned. Partitioning refers to the process of assigning part or all of the hard drive for use by the computer. It is a continuous section of sectors and is referred to by a drive letter, such as C: or D:. A hard drive can be partitioned into more than one partition and will then appear to have more than one physical hard drive. FDISK.EXE is a utility that is used to create and modify partitions and is one of the files that is found on the Windows Startup Disk.

The first partition that is created is the Primary partition and should be marked as active. This is the location of the DOS or Windows boot-up files. If there is more than one partition, the second and remaining partition is called an extended partition. An extended partition contains one or more logical partitions, which also have drive letters associated with them. An extended partition never has a drive letter associated with it. Also, a logical partition can only exist within an extended partition.

FDISK has four options, unless more than one physical hard drive exists in a machine, then a fifth option, to select the appropriate physical disk, also appears. FDISK allows you to create, delete, and display partitions as well as mark partitions as active. Marking a partition as active allows it to become bootable. One active partition is required for a Win9x installation.

> FDISK creates a start and an end to a section of hard drive space. At the beginning, the MBR, or Master Boot Record, is created. The MBR contains information about the beginning and end of the primary and extended partitions.
>
> The size of a partition affects the behavior of the File Allocation Table, or FAT. Generally, the larger the hard drive, the more wasted space there is on the drive.

Formatting

After a drive has been partitioned, a high level format is required and initiated for the FORMAT command. High level formatting is the only formatting a technician usually does. High level formatting:

a. Scans the surface of the drive and marks bad spots as bad sectors, so no data can be written to it.
b. Creates tracks, and sectors and CRC (error checking) areas.
c. Creates the File Allocation Table (FAT), which keeps track of location of files as they are placed onto the hard drive. The FAT has a limited number of entries.

Note: These days, the manufacturer of the drive low level formats the drive, so that today's technician will almost never need to do this. Low level formatting is the process of preparing the surface material of the hard disk to accept information. It creates the tracks and sectors that higher level processes look for. On older drives that needed a low level format, the low level formatting program is embedded into the hard drive controller interface. For exact instructions on how to perform this type of format, the manufacturer's Web site usually has the specific information.

Optimization

Because FAT has a finite limit to the amount of information that it can track, it creates limits on the number of sectors that it can track. Only high density floppy disks, because they have rather small capacities, have FATs that track all of the individual sectors. On the other hand, hard drives have a FAT that uses clusters, or allocation units, or groups of sectors, to track information.

With this in mind, a hard drive that is:

16Mb ? 128Mb uses 4 sectors per cluster

256MB   uses 8 sectors per cluster

512MB uses 16 sectors per cluster

1024MB (1 gig) uses 32 sectors per cluster (The 32 sector clusters are 16kb in size.)

2048MB (2 gig) used 64 sectors per cluster (The 64 sector clusters are 32kb in size.)

Because each allocation unit is considered used, even if all the clusters are not completely filled, there is a lot of wasted space with larger cluster size, and, as such, a lot of wasted space with larger hard drives.

Solution A: create multiple smaller partitions that will, in turn, use smaller clusters. For example a 528Mb partition has more wasted space, i.e. less available space on it, than a 512Mb partition.

Using the Windows 98 Boot Disk

This startup diskette is a bit different than a Win95 Startup Diskette in that it includes the option of booting with or without CD-ROM. Also, the Win98 FDISK program includes a message about working with large hard drives. This is because the Win98 FDISK program includes support for larger hard drives (i.e. includes support for FAT32).

Solution B: use FAT32, instead of FAT (FAT12 or FAT16, which refers to 12 bit or 16 bit FAT entries.) Beginning with the release of Windows 95B (OSR2), FAT32 became available. FAT32 includes support for larger hard drives, i.e., the cluster sizes are 4kb, instead of 64kb. If you save a 15KB file with FAT16, which uses 32KB cluster sizes, the entire 32KB cluster will be allocated. If you save the 15KB file with FAT32, using 4KB cluster sizes, then 4 4KB clusters would be allocated. This amounts to a savings of 16KB of disk space when compared to using FAT16.

Note: This also explains why one cannot boot a Win95B, (FAT32) system using a DOS boot floppy? It is not capable of seeing the FAT32 partition.

Starting a Windows 9x Installation

After the hard disk has been partitioned and formatted, then the computer can be booted using the Windows Startup Diskette and place the CD-ROM in the drive. The following notes outline the installation steps using a CD-ROM (preferred method) versus using a floppy disk install.

The program that performs the installation is called SETUP.EXE. SETUP.EXE is located either on Disk 1 or in the WIN95/WIN98 directory of the installation CD-Rom. There are various SETUP.EXE switches that you can use, such as Setup.exe /id.

Setup copies a very basic Windows system to the hard disk and then starts the Windows environment and welcomes you to the installation. The Setup Wizard informs you that it will perform the installation through three main steps: Gathering Information, Copying files to your computer, and Finishing the Installation. This is followed by the License Agreement, which you must accept.

Note: It is often a good idea to copy all the Win9x files from the Win9x directory on the CDROM to the hard drive of the machine. This way, if Windows needs these files later on, such as during a network configuration change perhaps, they are already on the hard drive.

- **Step 1**: Collecting Information about Your Computer

Choosing the Windows Installation Directory - Choose the Installation location, or use the suggested C:\Windows location. Note: Choose another directory if you wish to have both Windows 98 and Windows xp installed on the same computer. Setup also checks for adequate disk space.

Choosing the Setup Options ? There are four options to choose from:

Typical ? Installs the most popular features.

Portable ? Installs the most common features for laptops, including **PCMCIA** support, and Advanced Power Management (APM).

Compact ? Installed the bare minimum that Win9x needs to function.

Custom ? You can choose which options to install.

Entering User Information - This portion allows you to input user and company information.

Entering the Product Identification Number - Next, the Product Information Number (CD KEY from back of CD case or warranty card) must be entered.

Analyzing Your Computer and Setting Up Hardware - Setup can automatically look for hardware devices to install drivers for. Otherwise,

you can manually specify hardware as well. Hardware detection can take several minutes.

Setup can then look for drivers for special Setup options like a Network card, CD-ROM drive, or Sound/Midi/Video Capture Card.

The machine can sometimes lock up during this phase (after 10 minutes of waiting) and require a reboot. Rerun Setup and the process should resume where it left off.

Choosing E-Mail and Fax Software to Install - Next, the Get Connected screen, which enables user to decide to install software for connecting to the Internet (MSN), and using a fax modem (Microsoft Fax), or Email (Microsoft Mail).

Choosing Which Windows 98 Components You Want Installed - Then, you can choose which components to install, such as Accessibility Options, Accessories, Disk Tools, Multimedia, etc. When selecting which options to install, remember that a gray box with a check mark indicates that not all the components of that category are selected for install.

Network Configuration - The Networking configuration step is present next only if a network card (NIC) is present in the computer. This is followed by the computer identification screen which appears if Client for Microsoft Networks is installed.

Verifying Computer Settings - Finally, Setup displays a list of all the hardware detected on the system, which you can modify, if necessary.

Creating a Startup Disk - Following this, you have the opportunity to create a Startup Disk, if desired.

- **Step 2**: COPYING FILES TO YOUR COMPUTER

Setup copies all the information it gathered and detected in the previous step to your hard drive. You can cancel the installation with F3 or clicking the Exit button.

- **Step 3**: RESTARTING THE COMPUTER AND FINISHING SETUP

After the file copying is finished, Setup prompts you to reboot. This is followed by the words ?Getting ready to run Windows 98 for the first time??.

Setting Up Hardware and Software After Installation ? A screen is displayed next that reads? Windows 98 is now setting up your hardware and any Plug and Play devices you may have?

Setting the Date/Time Properties - Next, the Date and Time tab and the Time Zone tab are displayed. You can also elect to have the clock Automatically Adjust Clock for Daylight Saving Changes.

Setting Up Your Universal Inbox - Optional: If you elected to install MS Mail or MS Fax, setup will ask you to configure the email and fax services. Not covered in the exam.

Setting Up a Printer ? Windows 98 starts up the Add Printer Wizard to give you the opportunity to install a printer, if desired.

- **Final Installation Steps** - The final step in installation involves one final reboot.

Upgrading to Windows 98 from Windows 95

This upgrade is usually an easy process. Windows 98 enhancements include better Internet

Explorer integration, support for newer hardware, such as USB, AGP and DVD, FAT32 support, etc.

Starting the Upgrade - Once again the SETUP.EXE program, which is located on the Windows 98 CDROM can be used to perform the installation.

Preparing to Run Windows 98 Setup - Setup checks for adequate disk space. Remember to shut down all nonessential programs during the upgrade to avoid any potential conflicts.

Collecting Information about Your Computer ? A wizard will take you through the License Agreement, which you must accept. Next, the Product Key (CD KEY from back of CD case or warranty card) must be entered. This key is 25 characters long. Then, you must choose the Installation location, or use the suggested C:\Windows location. Setup again checks for adequate disk space.

Next, you must choose from four options what type of installation you wish to perform:

Typical ? Installs the most popular features.

Portable ? Installs the most common features for laptops, including PCMCIA support, and Advanced Power Management (APM).

Compact ? Installed the bare minimum that Win9x needs to function.

Custom ? You can choose which options to install. This option is usually best for a technician to select.

Then, the computer name, workgroup and description of the machine should be input. Setup confirms the keyboard layout, regional and country settings that you are using. The option to create a startup disk is available next.

Copying Windows 98 Files to Your Computer - Setup copies all the information it gathered from the CD to the hard drive.

Restarting Your Computer - After the file copying is finished, Setup prompts you to reboot and will automatically reboot after a 15-second delay.

Setting Up Hardware and Finalizing Settings ? Just like a Windows 95 install, Windows 98 next sets up detected hardware and any Plug and Play devices. Although Windows 98 will generally use the Windows 95 drivers previously in use, it is a good idea to check the web for updated Windows 98 specific drivers. Finally the machine will need to be rebooted again.

Windows 98 Setup ? During this phase, a screen is displayed that states that the Control Panel, Programs on the Start Menu, Windows Help, MS-DOS Program Settings, Tuning up Application Start and System **Configuration is being configured**. After this step, installation is finished.

Configuring Windows 9x Software

Although the Windows 9x registry holds most of the needed configuration information and replaces the INI files from Windows 3.x, many system files, such as the autoexec.bat and config.sys files, are present simply for compatibility with older hardware and software.

Important System Files

Examining the Windows 9x Boot Process

1. System self-checks and enumerates hardware resources ? The POST is executed from commands written to the motherboard.

2. MBR loads and finds the boot sector ? The master boot record is located on the first hard drive, loaded into memory, finds the bootable partition and searches for the boot sector of that partition. From there C:\IO.SYS is located.

3. IO.SYS loads into memory and starts the processor in real mode ? Real mode is simply a method of accessing the processor in 16-bit mode. The IO.SYS provides basic file system access to allow the rest of the boot files to be located. It accesses the MSDOS.SYS, LOGO.SYS, DRVSPACE.BIN files. It also loads the registry file SYSTEM.DAT into memory and selects a hardware profile. Lastly, it processes the CONFIG.SYS and AUTOEXEC.BAT files, if they are present.

4. WIN.COM loads and transfers the processor to protected mode ? Then WIN.COM loads various drivers as instructed by the registry and examines the SYSTEM.INI and WIN.INI files for more configuration information. The processor is then transferred into 32-bit protected mode.

5. Virtual device drivers, the Windows kernel, and the GDI load ? 32-bit virtual device are loaded. The Windows kernel, which controls access to the processor from Windows 9.x is loaded into memory and the GDI, or graphical display interface loads.

6. The Explorer shell loads and the user is finally presented with the Desktop ? Explorer, which manages the graphical interface. (the toolbar, the Desktop, Start menu), is loaded last. Network connections, if present, are restored and programs in the STARTUP folders are run.

Listing the Important Files

Windows 9.x system files are hidden from the user by default. To makes these files visible, open Windows Explorer, select View, Folder Options and click on the View tab. Scroll down until you can select Show All Files, deselect Hide File Extensions for Known File Types and then click OK.

Startup Files

An asterisk next to the file indicates that it is required to boot Windows 95.

- MSDOS.SYS* - primarily handles disk I/O
- EMM386.EXE ? provides the operating system with a mechanism to see additional memory, or upper memory and the upper memory blocks (UMB)
- HIMEM.SYS ? used to access upper memory
- IO.SYS* - allows the operating system and programs to interact between system hardware and the system BIOS. Includes built-in drivers for common hardware devices, such as printer ports, serial ports, floppy drives, hard drivers, etc.
- WIN.INI ? sets values for 16-bit Windows 3.x applications and is mostly replaced by the Registry for windows 9.x 32-bit applications.
- WIN.COM* - Initiates the Windows 9.x protected load phase.
- SYSTEM.INI ? Used to configure 16-bit DOS and Windows apps mostly in DOS and Windows 3.x.
- COMMAND.COM ? This ?DOS Shell?, or ?command interpreter? provides the command line interface, usually the C:\> prompt.
- CONFIG.SYS ? loads device drivers, such as memory managers and DOS peripheral drivers
- AUTOEXEC.BAT ? used to run particular programs at startup

Startup Files Configuration Tools

In order to edit the INI files, you can use NOTEPAD or other text editor, or you can use tools provided with Windows 9.x for this purpose.

SYSEDIT Click Start, Run and type SYSEDIT (short for SYSEDIT.EXE). From here, you can analyze and edit the PROTOCOL.INI, SYSTEM.INI, WIN.INI, CONFIG.SYS and AUTOEXEC.BAT files.

MSCONFIG Click Start, Run and type MSCONFIG (short for MSCONFIG.EXE). This System Configuration Utility is only included with Windows 98, not Windows 95. This editor allows you to use the mouse to configure Windows 98 elements by choosing from 6 tabs; General, Config.sys, Autoexec.bat, System.ini, Win.ini and Startup.

How Windows Works

How To Create A Boot Disk

WHAT IS A BOOT DISK

A boot disk will allow you to boot off of a diskette instead of your hard drive. This diskette can be used to fix issues which may arise during the lifetime of your computer as well as can be used to help load MS-DOS games or games you may not be able to run from Windows or MS-DOS because of high memory requirements.

After you have created a boot diskette following any of the below instructions it is highly recommended that you write protect the diskette to prevent possible computer virus infection.

CREATING A WINDOWS 98/ME BOOT DISK

An excellent feature of Windows 98/ME is its boot diskette. Using Windows to create you a Windows 98 boot diskette will give you all the needed files as well as CD-ROM support. To create a Windows 98 boot diskette click Start / Settings / Control Panel / double click the Add Remove programs icon / click the Startup Disk and create disk.

Alternatively to create a Windows 98 boot diskette manually from Windows click Start / Shutdown and choose the option to restart the computer in a MS-DOS prompt. At the prompt type:

cd\windows\command

Once at the C:\DOS directory skip to copying files.

COPYING ADDITIONAL FILES

NOTE: *When making a boot disk if you are running "Stacker" or some kind of a DoubleSpace or drive swapper program, this could not work*

Once you are in DOS and at the correct directory as instructed in the above sections by operating system you are ready to create your bootable diskette. Place a diskette which does not contain any information (it will be erased).

At the prompt if you have MS-DOS 6.2 / Windows 3.x / Windows 95 / Windows 98 type:

FORMAT A:/S

If you have MS-DOS 5.0 type using double density 5.25" diskettes type:

FORMAT A: /360 /S

If you have MS- DOS 3.11 through 4.0 type using double density 5.25" diskettes type:

FORMAT A: /4 /S

Once the diskette has been formatted and the system has been transferred you should be returned to your original directory. In this directory type:

copy format*.* a: [PRESS ENTER]
copy fdisk*.* a: [PRESS ENTER]
copy mscdex*.* a: [PRESS ENTER]
copy sys*.* a: [PRESS ENTER]
copy edit*.* a: [PRESS ENTER]
copy qbasic*.* a: [PRESS ENTER] (*Win 95/98 users skip this line*)
copy debug*.* a: [PRESS ENTER]
copy himem*.* a: [PRESS ENTER]
copy emm386*.* a: [PRESS ENTER]

If you are planning to use this diskette as a diskette to load games or you feel that you need mouse support you will need to copy the mouse driver onto the boot diskette. The MS-DOS mouse driver is generally mouse.com / mouse.sys. Locate this file and copy it to your bootable diskette.

For CD-ROM support visit our CD-ROM drivers page for information on loading your CD-ROM driver.

Once you have copied the above files create an autoexec.bat and a config.sys get to the floppy drive by typing **A:** once at the floppy drive type:

copy con autoexec.bat [PRESS ENTER]
@echo off [PRESS ENTER]
LH A:\MSCDEX.EXE /D:CDROM [PRESS ENTER] (this line is used for your CD-ROM drive)**.**
LH A:\MOUSE.* [PRESS ENTER] (skip line if you did not copy mouse file, the * is either sys / com).

Press and hold CTRL + Z - this should return ^Z once this is displayed press enter to copy the file.

copy con config.sys [PRESS ENTER]
device=a:\himem.sys

dos=high,umb
device=a:\emm386.exe noems
files=30
buffers=20
devicehigh=a:\oakcdrom.sys /d:CDROM (this line is used for your CD-ROM drive).

Press and hold CTRL + Z - this should return ^Z once this is displayed press enter to copy the file.

Congratulations after completing the above steps you should now have a bootable floppy diskette.

CREATING A WINDOWS NT BOOT DISK

To create a boot diskette you must have access to the i386 directory located on your Windows NT CD or possibly your Hard disk drive. Format the floppy diskette you wish to make a bootable Windows NT boot disk using the Windows NT machine. Copy boot.ini, ntdetect.com and ntldr to the floppy diskette just formatted. If you are using any SCSI devices which you need access to you will also need to load these drivers onto the diskette

CREATING A WINDOWS 2000 BOOT DISK

To create a Windows 2000 Professional bootable diskette you will need four 1.44MB diskettes and the Windows 2000 Professional CD. Click Start / Run / browse to the CD-ROM drive. Open the "BOOTDISK" folder and double-click makeboot.exe and click ok to launch the program to create the diskette.

CREATING A WINDOWS XP BOOT DISK

Microsoft is beginning to phase out bootable floppy diskettes in favor of bootable CD discs and has not included a method of easily creating a bootable floppy diskette in Windows or from the CD.

Using Fdisk Simulation

This is the first window that is displayed when opening FDISK. However this screen will not be displayed in Windows NT / Windows 95 OSR1 or lower operating systems. This screen is asking if you wish to enable the new FAT which is FAT32.

Your computer has a disk larger than 512 MB. This version of Windows includes improved support for large disks, resulting in more efficient use of disk space on large drives, and allowing disks over 2 GB to be formatted as a single drive.

IMPORTANT: If you enable large disk support and create any new drives on this disk, you will not be able to access the new drive(s) using other operating systems, including some versions of Windows 95 and Windows NT, as well as earlier versions of Windows and MS-DOS. In addition, disk utilities that were not designed explicitly for the FAT32 file system will not be able to work with this disk. If you need to access this disk with other operating systems or older disk utilities, do not enable large drive support.

Do you wish to enable large disk support (Y/N) . . . . ? [Y]

In this screen you will have two options to choose from, either Y for Yes to enable FAT32 when creating partitions. or N for No do not enable FAT32 and instead use FAT16.

The following screen is displayed if one or more of the hard drives within your computer is using NTFS. This screen will usually only be displayed if you are dual booting your computer with Windows 95 OSR2 / Windows 98 and Windows NT.

Your computer has NTFS partitions which may require large drive support. If you are using another operating system, such as Windows NT, which supports large drives you should enable treating these partitions as large. NOTE: If you answer Y and the partition display looks incorrect or a hang or crash occurs do nothing, run FDISK again, and answer N to this question.

Should NTFS partitions on all drives be treated as large (Y/N)? [Y]

Within this screen you have two options to choose from either Y for Yes to enable the drives to be treated as large. Or to not treat the drives as being large.

This is the main screen used during the time running FDISK. This would also be the first screen if your computer operating system does not support FAT32. From this window you will have the capability of setting up or removing partitions from your hard drive(s)

```
Microsoft Windows 98
Fixed Disk Setup Program
(C) Copyright Microsoft Corp. 1983 - 1998

FDISK Options

Current Fixed disk drive: 1

Choose one of the following:

1. Create DOS partition or Logical DOS Drive

2.  Set active partition
3.  Delete partition or Logical DOS Drive
4.  Display partition information
5.  Change current fixed disk drive

Enter choice:  [ 1 ]

Press Esc to exit FDISK
```

In this screen you will have the following options:

1. **Creating DOS partition or Logical DOS Drive -** Choose this option if you need to create a partition on your computer hard drive. If you are planning to recreate your partition you will first need to choose 3 to delete the partition.
2. **Set active partition -** If you have created multiple partitions on your hard drive choose this option to specify which partition you want to be the active partition. This is used to specify which partition the computer is going to look at to boot from. If this partition does not have an operating system however you will receive "Non system disk when booting up your computer." When choosing option 2 to set active partition you will see a listing of the current partitions choose the number representing the partition you wish to set active, the default is one.
3. **Delete partition or Logical DOS Drive -** This option is used if you need to delete partition(s) from the computers hard disk drive.
4. **Display partition information -** Use this option to look at the current Fixed disk drive partition information. If you have multiple drives you will need to choose option 5 before being able to display the other hard drives information.

5. **Change current fixed disk drive -** This option will only show if you have multiple drives within the computer. Using this option you will be able to toggle from one fixed drive to another.

Creating DOS partition Logical DOS partition

Using the options within this screen you will be able to create the various different partitions. These options will only allow you to create FAT16 and FAT32 partitions if supported and you pressed Y for Yes to enable large disk partitions. If you are attempting to create a NTFS partitions use Windows NT Disk Manager.

Create DOS Partition or Logical DOS Drive

Current fixed disk drive: 1

Choose one of the following:

1. Create Primary DOS Partition

2. Create Extended DOS Partition

3. Create Logical DOS Drive(s) in the Extended DOS Partition

Enter choice: [1]

1. **Create Primary DOS Partition -** Choosing this option you will be prompted to use maximum space. If you specify yes this will use up to 2 GB if creating FAT16 partitions or up to 32 GB if using FAT32. If you choose no you will be able to specify how large you would like the partition to be. NOTE you will need to create a primary partitions before being able to create an Extended or Logical DOS partitions.
2. **Create Extended DOS Partition -** If you are using FAT16 and have a 2 GB or higher hard drive or have only specified a small portion of the hard drive as the Primary partition use this option to create the Extended DOS partition(s) (other drive assignments). The Extended DOS partition will be used to hold the Logical DOS drives therefore use the maximum space left on the hard drive.

3. **Create Logical DOS Drive(s) in the Extended DOS Partition -** This option is used after you have created an Extended DOS partitions. Once the Extended DOS partition has been created you then can specify the sizes of other partitions you wish to create.

Example of what has been explained above:

A six GB hard disk drive needs to divide the hard drive into three partitions each using FAT 16.

Step 1. If prompted to use Large Disk support press N for no because Bob does not want FAT32 he would like FAT16.

Step 2. Choose option one to create a Primary partition. Once prompted to use maximum space press press **N** for no and specify 2 GB as the size of the primary partitions

Step 3. Once the primary DOS partition has been created choose option two to create an extended DOS partition. Use the maximum space which would be four GB because two GB has already been used for the Primary Partition.

Step 4. Create two logical DOS drives each being two GB.

Step 5. Reboot the computer and format each of the drives to allow them to be accessible.

Assuming Bob only had one hard drive doing the above Bob would have:

Drive C: Partition 1 (Primary) FAT 16

Drive D: Partition 2 (Extended/Logical) FAT16

Drive E: Partition 3 (Extended/Logical) FAT16

Note: Once a primary partition has been created please ensure that you set the partition as a active partition.

Deleting Partitions

Within this screen you will have the capability of deleting pre-existing DOS partitions. If you currently have no disk space available on your hard drive and wish to create additional partitions you must first use this screen to delete the partitions and then you

will be able to create partitions. NOTE if you delete partitions any information on those partitions will be erased and CANNOT be recovered.

```
Delete DOS Partition or Logical DOS Drive

Current fixed disk drive: 1

Choose one of the following:

1.  Delete Primary DOS partition

2.  Delete Extended DOS Partition

3.  Delete Logical DOS Drive(s) in the Extended DOS
Partition

4.  Delete Non-DOS Partition

Enter Choice:  [  ]



Press Esc to return to FDISK Options
```

1. **Delete Primary DOS partition -** Use this option to delete your main primary partition. However if you currently have any Extended / Logical DOS partitions you must delete these partitions before you will be able to delete the Primary DOS partition.
2. **Delete Extended DOS partition -** If you have your computer partitions into more then one drive, use this option to delete the extended dos partition(s). You must delete the Logical DOS Drive(s) before you can delete the Extended DOS partition.
3. **Delete Logical DOS Drive(s) in the Extended DOS Partition -** This option would be used first if you have extended DOS partitions and wish to delete the extended partitions.
4. **Delete Non-DOS Partition -** This option is usually used for partitions that either have been created by third party applications such as a DDO or other operating systems such as IBM Warp, Unix, as well as various other operating systems.

Example of what has been explained above:

We have created three partitions on one hard disk drive, however would like to delete them all.

Step 1. Delete the two logical DOS drive(s) in the Extended DOS partition with Option number three.

Step 2. Once the Logical DOS partitions have been deleted, choose option number two to delete the extended DOS partition.

Step 3. Choose option one to delete the Primary partition.

Step 4. Reboot the computer to allow above changes to take effect

Display Partition Information

Within this screen you will be able to see what is currently being used and how your computer hard drive is setup.

VIRUS

Display Partition Information

Current fixed disk drive: 1

Partition	Status	Type	Volume Label	Mbytes	System	Usage
C: 1	A	PRI DOS	WIN95	2047	FAT16	68%
			WIN95	2047	FAT16	68%
2		EXT	DOS	969		32%

Total disk space is 3016 Mbytes (1 Mbyte = 1048576 bytes)

The Extended DOS Partition contains Logical DOS Drives.
Do you want to display the logical drive information (Y/N) . . . . . . ? [Y] [Y]

Press Esc to return to FDISK Options

The above picture displays information about Extended DOS partitions if however you only have a Primary DOS partition your screen would only display the partition information and you will only have the option to Esc out of the screen. The following is what would displayed if Logical DOS drives existed within a Extended DOS partition. In the below example you can see one Extended DOS partition. However currently this partition is currently not formatted and this is why the system is UNKNOWN.

Display Logical DOS Drive Information

Drv	Volume	Label	Mbytes	System	Usage
D:			969	UNKNOWN	100%
D:			969	UNKNOWN	100%

Total Extended DOS Partition is 969 Mbytes (1 MByte = 1048576 bytes)

Press Esc to continue

Change Current Fixed Disk Drive

This option is only available if more then one Hard Drive is installed within the computer. In the below picture you can notice that we have 3 hard drives listed within this computer. Disk 1 has two partitions which are C: and D: the hard drives total space is 3 GB. Then on Disk 2 has three partitions E:, F:, and G: and the hard drive total space is 6668 or 7 GB. As you notice in the below picture there is also a Disk 3 this is actually a Iomega Zip Drive hooked up to the IDE controller on the Motherboard. Other media that is connected to the IDE controller within your computer will usually be displayed within FDISK this is not however a hard drive and does not need to be FDISKED this is why 95MB is free. An exception to this rule would be a CD-ROM.

Change Current Fixed Disk Drive

Disk	Drv	Mbytes	Free	Usage
1		3016		100%
	C:	2047		
	D:	969		
2		6668	1	100%
	E:	2047		
	F:	2047		
	G:	526		
3		95	95	%
1		3016		100%
	C:	2047		
	D:	969		
2		6668	1	100%
	E:	2047		
	F:	2047		
	G:	526		
3		95	95	%

(1 MByte = 1048576 bytes)
Enter Fixed Disk Drive Number (1-3) . . . . [1]

Press **Esc** to continue

Chapter 17

Managing & Troubleshooting Windows

OBJECTIVES

File Systems

Troubleshooting Windows 98

Installing New Hardware in Windows

Connecting to Internet in Windows 98

Windows Registry

File System

Physically, it is the hard disks and other media that provide the basic technology for storing data. Logically, however, it is the file system that provides the hierarchical structure of volumes and directories in which you store individual files and the organizational model that enables the system to locate data anywhere on a given disk or drive. File systems typically are an integrated part of an operating system (OS), and many of the newer OSes provide support for several file systems from which you can choose.

File Allocation Table

The most commonly used file systems today are based on a file allocation table (FAT), which keeps track of the data stored in each cluster on a disk. Several varieties of the FAT system exist, called FAT12, FAT16, and FAT32—all of which are differentiated by the number of digits used in the allocation table numbers. In other words, FAT16 uses 16-bit numbers to keep track of data clusters; FAT32 uses 32-bit numbers, and so on. The various FAT systems are used as follows:

- *FAT12*. Used on all volumes smaller than 16MB (for example, floppies)
- *FAT16*. Used on volumes from 16MB through 2GB by MS-DOS and most versions of Windows; Windows NT, Windows 2000, and Windows XP support FAT16 volumes as large as 4GB
- *FAT32*. Optionally used on volumes from 512MB through 2TB

Cluster

Also called *allocation unit*. A group of one or more sectors on a disk that forms a fundamental unit of storage to the operating system. Cluster, or allocation unit, size is determined by the operating system when the disk is formatted. Larger clusters generally offer faster system performance but waste disk space.

Boot Sector (VBS)

This sector contains data that describes the partitions by their starting and ending cylinder, head, and sector locations. The partition table also indicates to the ROM BIOS which of the partitions is bootable and, therefore, where to look for an operating system to load.

Root Directory

It is a directory in which the operating system stores information about the names and starting locations of the files on the disk.

Master Partition Boot Record

Information about each of the partitions on the disk is stored in a partition (or volume) boot record at the beginning of each partition. Additionally, a main table lists the partitions embedded in the master partition boot record.
The master partition boot record (or master boot sector) always is located in the first sector of the entire disk (cylinder 0, head 0, sector 1) and consists of the following structures:

- *Master partition table.* Contains a list of the partitions on the disk and the locations of their volume boot records. This table is very small and can contain only four entries at the most. Therefore, to accommodate more partitions, operating systems such as DOS can divide a single extended partition into multiple logical volumes.
- *Master boot code.* A small program that is executed by the system BIOS, the main function of which is to pass control of the system to the partition that has been designated as active (or bootable).

Volume Boot Records

The volume boot record is the first sector on any area of a drive addressed as a partition or volume (or logical DOS disk). On a floppy disk or removable cartridge (such as a Zip disk), for example, this sector is the first one on the disk because DOS recognizes the disk as a volume without the need for partitioning. On a hard disk, the volume boot record or sectors are located as the first sector within any disk area allocated as a non-extended partition, or any area recognizable as a logical volume.

Primary and Extended FAT Partitions

Most OSes are designed to support up to 24 partitions on a single hard disk drive (represented by the drive letters C: through Z:), but the partition table in the master partition boot record can have a maximum of only four entries. To resolve this discrepancy, the FDISK program enables you to create two types of FAT partitions: a primary partition and an extended partition. The first FAT partition you create on a disk should be the *primary*, which is listed in the master partition table and appears to the operating system as a single drive letter. An *extended* partition is listed in the master partition table like the primary, but it differs in that you can use its disk space to create multiple logical partitions, or volumes. You can create only one extended partition on a single drive, meaning that typically there will never be more than two entries in the master partition table devoted to FAT volumes.

FAT Disk Structures 1319

The logical volumes you create in the extended partition appear as separate drive letters to the operating system, but they are not listed in the master partition table. Logical volumes can't be active partitions and are therefore not bootable. You can create up to 23 volumes out of a single extended partition (assuming that you already have created a primary FAT partition, which brings the total to 24).

Clusters (Allocation Units)

The term *cluster or Allocation unit* is the smallest unit of the disk that the operating system can handle when it writes or reads a file. A cluster is equal to one or more 512-byte sectors, in a power of two. Although a cluster can be a single disk sector, it is usually more than one. Having more than one sector per cluster reduces the size and processing overhead of the FAT and enables the operating system to run faster because it has fewer individual units to manage. The trade off is in wasted disk space.

Default Floppy Disk Cluster (Allocation Unit) Sizes

Drive Type	Cluster (Allocation Unit) Size	Density
5 1/4-inch 360KB	2 sectors (1,024 bytes)	Low
5 1/4-inch 1.2MB	1 sector (512 bytes)	High
3 1/2-inch 720KB	2 sectors (1,024 bytes)	Low
3 1/2-inch 1.44MB	1 sector (512 bytes)	High
3 1/2-inch 2.88MB	2 sectors (1,024 bytes)	Extra

The Data Area

The data area of a partition is the place where the actual files are stored. It is located following the boot record, file allocation tables, and root directory. This is the area of the disk that is divided into clusters and managed by the FAT and root directory.

Diagnostic Read-and-Write Cylinder

The FDISK partitioning program always reserves the last cylinder of a hard disk for use as a special diagnostic read-and-write test cylinder. Because this cylinder is reserved, FDISK always reports fewer total cylinders than the drive manufacturer states are available. Operating systems do not use this cylinder for any normal purpose because it lies outside the partitioned area of the disk. The diagnostics area enables software such as a manufacturer-supplied diagnostics disk to perform read-and-write tests on a hard disk without corrupting any user data. Many of these programs also swap spare cylinders for damaged cylinders if damaged cylinders are detected during testing.

FAT32

FAT32 works just like the standard FAT; the only difference is that it uses numbers with more digits, so it can manage more clusters on a disk. Unlike VFAT, which is a Windows 9x innovation that uses existing file system structures, FAT32 is an enhancement of the FAT file system itself. In other words, whereas VFAT is implemented as part of the Windows 9x Virtual Machine Manager (Vmm.vxd), FAT32 is implemented by the FDISK program before the Windows GUI is even loaded. FAT32 was first included in the Windows 95 OEM Service Release 2 (OSR2) and is also part of Windows 98/Me, Windows 2000, and Windows XP operating systems. Because FAT32 partitions can have so many more clusters than FAT16 partitions, the clusters themselves can be smaller.

Converting FAT16 to FAT32

If you want to convert an existing FAT16 partition to FAT32, you must back up the data, destroy the FAT16 partition, create a new FAT32 partition, and then restore the data. Windows 98/Me, on the other hand, includes a FAT32 Conversion Wizard that enables you to migrate existing partitions in place. The wizard gathers the information needed to perform the conversion, informs you of the consequences of implementing FAT32, and attempts to prevent data loss and other problems. After you have selected the drive you want to convert, the wizard performs a scan for applications (such as disk utilities) that might not function properly on the converted partition. The wizard gives you the opportunity to remove these and warns you to back up the data on the partition before proceeding with the conversion. Even if you don’t use the Microsoft Backup utility the wizard offers, backing up your data is a strongly recommended precaution.

FAT File System Errors

File system errors can, of course, occur because of hardware problems, but you are more likely to see them result from software crashes and improper system handling. Turning off a system without shutting down Windows properly, for example, can result in errors that cause clusters to be incorrectly listed as in use when they are not. Some of the most common file system errors that occur on FAT partitions are described in the following sections.

Lost Clusters

Probably the most common file system error, lost clusters are clusters the FAT designates as being in use when they actually are not. Most often caused by an interruption of a file system process due to an application crash or a system shutdown, for example, the FAT entry of a lost cluster might contain a reference to a subsequent cluster. However, the FAT chain stemming from the directory entry has been broken somewhere along the line.

Lost clusters would appear in the file structure as shown in Table.

Directory

Name	Starting Cluster	Size
(no entry)	0	0

FAT16 File Allocation Table

FAT Cluster #	Value	Meaning
00002	0	First cluster available
...	...	...
00999	0	Cluster available
01000	1001	In use, points to next cluster
01001	1002	In use, points to next cluster
01002	1003	In use, points to next cluster
01003	FFFFh	End of file
01004	0	Cluster available
...	...	...
65526	0	Last cluster available

The operating system sees a valid chain of clusters in the FAT but no corresponding directory entry to back it up. Programs that are terminated before they can close their open files typically cause this. The operating system usually modifies the FAT chain as the file is written, and the final step when closing is to create a matching directory entry. If the system is interrupted before the file is closed (such as by shutting down the system improperly), lost clusters are the result. Disk repair programs check for lost clusters by tracing the FAT chain for each file and subdirectory in the partition and building a facsimile of the FAT in memory. After compiling a list of all the FAT entries that indicate properly allocated clusters, the program compares this facsimile with the actual FAT. Any entries denoting allocated clusters in the real FAT that do not appear in the facsimile are lost clusters because they are not part of a valid FAT chain. The utility typically gives you the opportunity to save the data in the lost clusters as a file before it changes the FAT entries to show them as unallocated clusters. If your system crashed or lost power while you were working with a word processor data file, for example, you might be able to retrieve text from the lost clusters in this way. When left un repaired, lost clusters are unavailable for use by the system, reducing the storage capacity of your drive. The typical choices you have for correcting lost clusters are to assign them a made-up name or zero out the FAT entries. If you assign them a name, you can at least look at the entries as a valid file and then delete the file if you find it useless. The CHKDSK and SCANDISK programs are designed to fix lost clusters by assigning them names starting with FILE0001.CHK. If more than one lost chain exists, sequential numbers following the first one are used. The lost clusters shown earlier could be corrected by CHKDSK or SCANDISK, as shown in Table .

Finding Lost Clusters

Directory		
Name	**Starting Cluster**	**Size**
FILE0001.CHK	1000	4

FAT16 File Allocation Table

FAT Cluster #	**Value**	**Meaning**
00002	0	First cluster available
...	...	...
00999	0	Cluster available
01000	1001	In use, points to next cluster
01001	1002	In use, points to next cluster
01002	1003	In use, points to next cluster
01003	FFFFh	End of file
01004	0	Cluster available
...	...	...
65526	0	Last cluster available

Cross-Linked Files

Cross-linked files occur when two directory entries improperly reference the same cluster in their Link to Start Cluster fields. The result is that each file uses the same FAT chain. Because the clusters can store data from only one file, working with one of the two files can inadvertently overwrite the other file's data. Cross-linked files would appear in the file structure as shown in Table.

Cross-Linked Files

Directory		
Name	**Starting Cluster**	**Size**
USCONST.TXT	1000	4
PLEDGE.TXT	1002	2

FAT16 File Allocation Table

FAT Cluster #	**Value**	**Meaning**
00002	0	First cluster available
...	...	...
00999	0	Cluster available
01000	1001	In use, points to next cluster
01001	1002	In use, points to next cluster
01002	1003	In use, points to next cluster
01003	FFFFh	End of file
01004	0	Cluster available
...	...	...
65526	0	Last cluster available

In this case, two files claim ownership of clusters 1002 and 1003. These files are said to be cross-linked on 1002. When a situation such as this arises, one of the files typically is valid and the other is corrupt, being that only one actual given set of data can occupy a given cluster. The normal repair is to copy both files involved to new names, which duplicates their data separately in another area of the disk, and then delete *all* the files that are cross-linked. Deleting them all is important because by deleting only one of them, the FAT chain is zeroed, which further damages the other entries.

Invalid Files or Directories

Sometimes the information in a directory entry for a file or subdirectory can be corrupted to the point at which the entry is not just erroneous (as in cross-linked files) but invalid. The entry might have a cluster or date reference that is invalid, or it might violate the rules for the entry format in some other way. In most cases, disk repair software can correct these problems, permitting access to the file.

FAT File System Utilities

The CHKDSK, RECOVER, and SCANDISK commands are the DOS damaged-disk recovery team. These commands are crude, and their actions sometimes are drastic, but at times they are all that is available or necessary.

The *RECOVER* Command

The DOS RECOVER command is designed to mark clusters as bad in the FAT when the clusters can't be read properly. When the system can't read a file because of a problem with a sector on the disk going bad, the RECOVER command can mark the FAT so another file does not use those clusters. When used improperly, this program is highly dangerous. The RECOVER utility has not been included in DOS since version 5 and is not supplied in Windows 9x because its functionality has been replaced by SCANDISK.

SCANDISK

You should check your FAT partitions regularly for the problems discussed in this chapter and any other difficulties that might arise. By far an easier and more effective solution for disk diagnosis and repair than CHKDSK and RECOVER is the SCANDISK utility, included with DOS 6 and higher versions, as well as with Windows 9x. This program is more thorough and comprehensive than CHKDSK or RECOVER and can perform the functions of both of them—and a great deal more. Windows 95 OSR2 and Windows 98 include SCANDISK versions that support FAT32 partitions.

Disk Defragmentation

The entire premise of the FAT file systems is based on the storage of data in clusters that can be located anywhere on the disk. This enables the computer to store a file of nearly any size at any time. The process of following a FAT chain to locate all the clusters holding the data for a particular file can force the hard disk drive to access many locations on the disk. Because of the physical work involved in moving the disk drive heads, reading a file that is heavily fragmented in this way is slower than reading one that is stored on consecutive clusters.

Third-Party Programs

When you get a Sector not found error reading drive C:, the best course of action is to use one of the third-party disk-repair utilities on the market, rather than DOS's RECOVER or even SCANDISK. The Norton Utilities by Symantec (also included in Norton System Works) stands as perhaps the premier data recovery package on the market today. This package is comprehensive and automatically repairs most types of disk problems.
Norton Utilities and Norton System Works Programs such as Norton Disk Doctor can perform much more detailed repairs with a greater amount of safety. Disk Doctor preserves as much of the data in the file as possible and can mark the FAT so the bad sectors or clusters of the disk are not used again. These programs also save Undo information, enabling you to reverse any data recovery operation.

Disk Doctor is part of Symantec's Norton Utilities package, which includes a great many other useful tools. For example, Norton Utilities has an excellent sector editor that enables you to view and edit any part of a disk, including the master and volume boot records, FATs, and other areas that fall outside the disk's normal data area. Currently, no other program is as comprehensive or as capable of editing disks at the sector level. The disk editor included with Norton Utilities can give the professional PC troubleshooter or repairperson the ability to work directly with any sector on the disk, but this does require extensive knowledge of sector formats and disk structures. The documentation with the package is excellent and can be very helpful if you are learning data recovery on your own.

NTFS

NTFS is the native file system of Windows NT, Windows 2000, and Windows XP. Windows 2000 and Windows XP use an enhanced version of NTFS called NTFS 5 or NTFS 2000; Windows NT 4.0 must have Service Pack 4 or above installed to be capable of accessing an NTFS 5/NTFS 2000 disk. Although NT/2000/XP support FAT partitions (and Windows 2000/XP even support FAT32), NTFS provides many advantages over FAT, including long filenames, support for larger files and partitions, extended attributes,

and increased security. NTFS, like all of Windows NT, was newly designed from the ground up. Backward compatibility with previous Microsoft operating systems was not a concern because the developers were intent on creating an entirely new 32-bit platform. As a result, no operating systems other than Windows NT and Windows 2000/XP, which are based on Windows NT, can read NTFS partitions. This is an important consideration. All the other Windows operating systems are fundamentally based on DOS and leave you the alternative of booting to a DOS prompt in special situations. Whether you have to run a special DOS utility to configure a piece of hardware or perform an emergency disk repair, the DOS option is always there. Windows NT/2000/XP are not DOS based. You can open a window that provides a DOS-like command prompt in Windows NT/2000/XP, but this is actually a DOS emulation—not a true shell. You can't boot the system to a command prompt, avoiding the GUI, as you can with Windows 9x unless you install the Windows 2000 or Windows XP Recovery Console. As the name implies, the Recovery Console is designed for system repair and maintenance procedures, not for routine use. You can, of course, bypass NT/2000/XP entirely by starting the system with a DOS boot disk, but if you have NTFS partitions on your drives, you can't access them from the DOS prompt. NTFS supports filenames of up to 255 characters, using spaces, multiple periods, and any other standard characters except the following: *?/\;<>|. Because NTFS file offsets are 64 bits long, files and partitions can be truly enormous: up to 16EB (exabytes) in size (1EB = 264 bytes = 17,179,869,184TB)! To give you an idea of just how much data this is, it is estimated that all the words ever spoken by humans throughout history would occupy 5EB of storage.

Since Windows NT 3.51, NTFS has also supported compression on a file-by-file basis through each file's properties sheet. No third-party program, such as WinZip or PKZip, is needed to compress or decompress files stored on an NTFS drive.

NTFS Architecture

Although NTFS partitions are very different from FAT partitions internally, they do comply with the extra-partitional disk structures described earlier in this chapter. NTFS partitions are listed in the master partition table of a disk drive's master boot record, just like FAT partitions, and they have a volume boot record as well, although it is formatted somewhat differently. When a volume is formatted with NTFS, system files are created in the root directory of the NTFS volume. These system files can be stored at any physical location on the NTFS volume. This means that damage to any specific location on the disk will probably not render the entire partition inaccessible. Typically, 10 NTFS system files are created when you format an NTFS volume. Table shows the names and descriptions for these files.

NTFS System Files

Filename	Meaning	Description
$mft	Master file table (MFT)	Contains a record for every file on the NTFS volume in its Data attribute
$mftmirr	Master file table2 (MFT2)	Mirror of the MFT used for recoverability purposes
$badclus	Bad cluster file	Contains all the bad clusters on the volume
$bitmap	Cluster allocation bitmap	Contains the bitmap for the entire volume, showing which clusters are used
$boot	Boot file	Contains the volume's bootstrap if the volume is bootable
$attrdef	Attribute definitions table	Contains the definition of all system- and user-defined attributes on the volume
$logfile	Log file	Logs file transactions, used for recoverability purposes
$quota	Quota table	Table used to indicate disk quota usage for each user on a volume, used in NTFS 5
$upcase	Upcase table	Table used for converting uppercase and lowercase characters to the matching uppercase Unicode characters
$volume	Volume	Contains volume information, such as volume name and version

An NTFS partition is based on a structure called the master file table (MFT). The MFT concept expands on that of the FAT. Instead of using a table of cluster references, the MFT contains much more detailed information about the files and directories in the partition. In some cases, it even contains the files and directories themselves. The first record in the MFT is called the descriptor, which contains information about the MFT itself. The volume boot record for an NTFS partition contains a reference that points to the location of this descriptor record. The second record in the MFT is a mirror copy of the descriptor, which provides fault tolerance, should the first copy be damaged. The third record is the log file record. All NTFS transactions are logged to a file that can be used to restore data in the event of a disk problem. The bulk of the MFT consists of records for the files and directories stored on the partition. NTFS files take the form of objects that have both user- and system-defined attributes. Attributes on NTFS partitions are more comprehensive than the few simple flags used on FAT partitions. All the information on an NTFS file is stored as attributes of that file. In fact, even the file data itself is an attribute. Unlike FAT files, the attributes of NTFS files are part of the file itself; they are not listed separately in a directory entry. Directories exist as MFT records as well, but they consist mainly of indexes listing the files in the directory; they do not contain the size, date, time, and other information about the individual files. Thus, an NTFS drive's MFT is much more than a cluster list, like a FAT; it is actually the primary data storage structure on the partition. If a file or directory is relatively small (less than approximately 1,500 bytes), the entire file or directory might even be stored in the MFT. For larger amounts of storage, the MFT record for a file or directory contains pointers to external clusters in the partition. These external clusters are called *extents*. All the records in the MFT, including the descriptors and the log file, are capable of using extents for storage of additional attributes. The attributes of a file that are part of the MFT record are called *resident* attributes. Those stored in extents are called *nonresident* attributes.

NTFS 5 (NTFS 2000)

Along with Windows 2000 came a new variation of NTFS called NTFS 5 (also called NTFS 2000); NTFS 5 is also used by Windows XP. This update of the NT file system includes several new features that are exploited and even required by Windows 2000. Because of this, when you install Windows 2000, any existing NTFS volumes automatically are upgraded to NTFS 5 (there is no way to override). If you also run Windows NT versions earlier than Windows NT 4 Service Pack 4 (SP4), NT 4 is no longer capable of accessing the NTFS 5 volumes. If you want to run both NT 4 and Windows 2000 on the same system (as in a dual-boot configuration), you must upgrade NT 4 by installing Service Pack 4 or later. An updated NTFS.SYS driver in Service Pack 4 enables NT 4 to read from and write to NTFS 5 volumes. New features of the NTFS 5 file system include

- *Disk quotas.* System administrators can limit the amount of disk space users can consume on a per-volume basis. The three quota levels are Off, Tracking, and Enforced.
- *Encryption.* NTFS 5 can automatically encrypt and decrypt files as they are read from and written to the disk.
- *Reparse points.* Programs can trap open operations against objects in the file system and run their own code before returning file data. This can be used to extend file system features such as mount points, which you can use to redirect data read and written from a folder to another volume or physical disk.
- *Sparse files.* This feature enables programs to create very large files as placeholders but to consume disk space by adding to the files only as necessary.
- *USN (Update Sequence Number) Journal.* Provides a log of all changes made to files on the volume. Because these features—especially the USN Journal—are required for Windows 2000 to run, a Win2000 domain controller must use an NTFS 5 partition as the system volume.

NTFS Compatibility

Although NTFS partitions are not directly accessible by DOS and other operating systems, Windows NT/2000 is designed for network use, so other operating systems are expected to be capable of accessing NTFS files via the network. For this reason, NTFS continues to support the standard DOS file attributes and the 8.3 FAT naming convention. One of the main reasons for using NTFS is the security it provides for its files and directories. NTFS security attributes are called *permissions* and are designed to enable system administrators to control access to files and directories by granting specific rights to users and groups. This is a much more granular approach than the FAT file system attributes, which apply to all users. However, you can still set the FAT-style attributes on NTFS files using the standard Windows NT/2000 file management tools, including Windows NT/2000 Explorer and even the command-prompt ATTRIB command. When you copy FAT files to an NTFS drive over the network, the FAT-style attributes remain in place until you explicitly remove them. This can be an important consideration because the FAT-style attributes take precedence over the NTFS

permissions. A Windows NT/2000 user, for example, can't delete a file on an NTFS drive that is flagged with the FAT read-only attribute, even if that user has NTFS permissions that grant her full access. To enable DOS and 16-bit Windows systems to access files on NTFS partitions over a network, the file system maintains an 8.3 alias name for every file and directory on the partition. The algorithm for deriving the alias from the long filename is the same as that used by Windows 95's VFAT. Windows NT/2000 also provides its FAT partitions with the same type of long filename support used by VFAT, allocating additional directory entries to store the long filenames as needed.

Creating NTFS Drives

NTFS is for use on hard disk drives. You can't create an NTFS floppy disk (although you can format removable media, such as Iomega Zip and Jaz cartridges to use NTFS). Three basic ways to create an NTFS disk partition are as follows:

- Create a new NTFS volume out of unpartitioned disk space during the Windows NT/2000 installation process or after the installation with the Disk Administrator utility.
- Format an existing partition to NTFS (destroying its data in the process), using the Windows NT Format dialog box (accessible from Windows NT Explorer or Disk Administrator) or the FORMAT command (with the /fs:ntfs switch) from the command prompt.
- Convert an existing FAT partition to NTFS (preserving its data) during the Windows NT/2000 installation process or after the installation with the command-line CONVERT utility.

NTFS Tools

Because it uses a fundamentally different architecture, virtually none of the troubleshooting techniques outlined earlier in this chapter are applicable when dealing with NTFS partitions, nor can the disk utilities intended for use on FAT partitions address them. Windows NT has a rudimentary capability to check a disk for file system errors and bad sectors with its own version of CHKDSK, but apart from that, the operating system contains no other disk repair or defragmentation utilities. Windows 2000 and Windows XP include a commmand-line and GUI version of CHKDSK and also include a defragmenting tool that is run from the Windows Explorer GUI. The NTFS file system, however, does have its own automatic disk repair capabilities. In addition to Windows NT/2000/XP's fault-tolerance features, such as disk mirroring (maintaining the same data on two separate drives) and disk striping with parity (splitting data across several drives with parity information for data reconstruction), the OS also uses a disk error recovery technique called *cluster remapping*. With cluster remapping, when Windows (NT/2000/XP) detects a bad sector on an NTFS partition, it automatically remaps the data

in that cluster to another cluster. If the drive is part of a fault-tolerant drive array, any lost data is reconstructed from the duplicate data on the other drives. Despite these features, however, there is still a real need for disk repair and defragmentation utilities for Windows NT/2000/XP. They have been a long time in coming, but finally third-party utilities now are available that can repair and defragment NTFS drives. One I recommend is Norton Utilities 2001 by Symantec, which works with Windows 98, Windows Me, Windows NT 4.0, and Windows 2000. An updated version is necessary for Windows XP.

Common Drive Error Messages and Solutions

Several common error messages are associated with problems in the file system or drives. This section covers the common ones, listing their causes and possible solutions.
Some of the most common file system errors are as follows:

- Missing Operating System
- NO ROM BASIC - SYSTEM HALTED
- Boot error Press F1 to retry
- Invalid drive specification
- Invalid Media Type
- Hard Disk Controller Failure

These normally occur when booting the system or when trying to log in to or access a drive. The following sections describe each of these errors and offer possible solutions.

Missing Operating System

This error indicates problems in the master boot record or partition table entries. The partition table entries might be pointing to a sector that is not the actual beginning of a partition. This can also be caused by invalid BIOS settings, in some cases resulting from a dead or dying battery. In fact, in my experience, corrupt or improper BIOS settings—such as incorrect geometry (cylinder/head/sector) or LBA mode settings—are the number-one cause of this problem. Another cause can be virus damage to the MBR. This error also can occur if no active partition is defined in the partition table. The normal solution is to correct the invalid BIOS settings. The BIOS settings for drive parameters and LBA translation must be set to the same values as when the drive was partitioned and formatted to read the drive correctly. If the MBR is damaged or virus infected, you can try FDISK/MBR to repair it. Other types of damage require more sophisticated use of a disk editor utility or repartitioning and reformatting the drive to start over.

NO ROM BASIC - SYSTEM HALTED

This error is generated by the AMI BIOS when the boot sector or master boot record of the boot drive is damaged or missing. This error also can occur if the boot device has been improperly configured or is not configured at all in the BIOS. In this case, data in the partition might be valid and undamaged, but no bootable partition exists. IBM systems in this situation used to drop into a built-in BIOS version of BASIC, but most non-IBM BIOS manufacturers did not license this code from Microsoft. So, instead of dropping into BASIC, they would display this cryptic message. Because the most common cause of this type of error is a failure to set at least one partition as active (bootable), the normal solution is to run FDISK and set the primary partition as active. If this is not the problem, the solution is to repair the damaged MBR or correct the improper BIOS settings. Other possible problems include corrupted or missing drive parameters in the BIOS Setup.

Boot Error Press F1 to Retry

This error is generated by the Phoenix BIOS when the hard disk is missing a master boot record or boot sector or when there is a problem accessing the boot drive. This has the same meaning as NO ROM BASIC does on an AMI BIOS. The most common cause of this message is having no partitions defined as active (bootable). See the section on NO ROM BASIC earlier in the chapter for more information. General File System Troubleshooting for MS-DOS, Windows 9x, and Windows Me

Invalid Drive Specification

This error occurs when you attempt to log in to a drive that has not been partitioned or for which the partition table entry has been damaged or is incorrect. Use FDISK to partition the drive or to check out the existing partitions. If they are damaged, you probably should use a data recovery utility such as the Norton Utilities to correct the problem. Such correction might require manual editing of the partition table with the DISKEDIT program included with the Norton Utilities. Another solution is to repartition the drive from scratch, but this causes any existing data on the drive to be overwritten.

Invalid Media Type

This indicates the partition table is valid, but the volume boot sector, directory, or file allocation tables are corrupt, damaged, or not yet initialized. For example, this is the standard error you would receive if you tried to access a drive that had been partitioned but not yet formatted. The FORMAT command is what creates the volume boot record (VBR), file allocation tables, and directories on the disk. The repair usually involves using a data recovery utility, such as the Norton Utilities Disk Doctor, or redoing the high-level format on the drive. Because high-level formatting does not actually destroy the data, one technique to recover is to high-level format (OS Format) the volume and

then immediately unformat it using the unformat utility included with the Norton Utilities.

Hard Disk Controller Failure

This message indicates the hard disk controller has failed, the hard disk controller is not set up properly in the BIOS, or the controller can't communicate with the attached drives, such as with cable problems. The solution is to check out the drive installation and ensure that the cables to the drive are properly installed, the drive is receiving power, it is spinning, and the BIOS Setup definitions are correct. If all these are correct, the drive, cable, or controller might be physically damaged. Replace them with known-good spares one at a time until the problem is solved.

General File System Troubleshooting for MS-DOS, Windows 9x, and Windows Me

Here are some general procedures to follow for troubleshooting drive access, file system, or boot problems:

1. Start the system using a Windows startup disk, or any bootable MS-DOS disk that contains FDISK.EXE, FORMAT.COM, SYS.COM, and SCANDISK.EXE (Windows 95B or later versions preferred).

2. If your system can't boot from the floppy, you might have more serious problems with your hardware. Check the floppy drive and the motherboard for proper installation and configuration. On some systems, the BIOS configuration doesn't list the floppy as a boot device or puts it after the hard disk. Reset the BIOS configuration to make the floppy disk the first boot device if necessary and restart your computer.

3. Run FDISK from the Windows startup disk. Select option 4 (Display partition information).

4. If the partitions are listed, make sure that the bootable partition (normally the primary partition) is defined as active (look for an uppercase *A* in the Status column).

5. If no partitions are listed and you do not want to recover any of the data existing on the drive now, use FDISK to create new partitions, and then use FORMAT to format the partitions. This overwrites any previously existing data on the drive.

6. If you want to recover the data on the drive and no partitions are being shown, you must use a data recovery program, such as the Norton Utilities by Symantec or Lost and Found by PowerQuest, to recover the data.

7. If all the partitions appear in FDISK.EXE and one is defined as active, run the SYS command as follows to restore the system files to the hard disk:

SYS C:

8. For this to work properly, it is important that the disk you boot from be a startup disk from the same operating system (or version of Windows) you have on your hard disk.

9. You should receive the message System Transferred if the command works properly. Remove the disk from drive A:, and restart the system. If you still have the same error as before after you restart your computer, your drive might be improperly configured or damaged.
10. Run SCANDISK from the Windows startup disk or an aftermarket data-recovery utility, such as the Norton Utilities, to check for problems with the hard disk.
11. Using SCANDISK, perform a surface scan. If SCANDISK reports any physically damaged sectors on the hard disk, the drive might need to be replaced.
General File System Troubleshooting for

Windows 2000/XP

The process for file system troubleshooting with Windows 2000/XP is similar to that used for
Windows 9x. The major difference is the use of the Windows 2000/XP Recovery Console, which is clarified here:

- If the Recovery Console was added to the boot menu, start the system normally, log in as Administrator if prompted, and select the Recovery Console.
- If the Recovery Console was not previously added to the boot menu, start the system using the Windows CD-ROM or the Windows Setup disks. Select Repair from the Welcome to Setup menu, and then press C to start the Recovery Console when prompted. If your system can't boot from CD-ROM or the floppy, you might have more serious problems with your hardware. Check your drives, your BIOS configuration, and the motherboard for proper installation and configuration. Set the floppy disk as the first boot device and the CD-ROM as the second boot device and restart the system.

After you start the Recovery Console do the following:

1. Type **HELP** for a list of Recovery Console commands and assistance.
2. Run DISKPART to examine your disk partitions.
3. If the partitions are listed, make sure that the bootable partition (normally the primary partition)
is defined as active.
4. If no partitions are listed and you do not want to recover any of the data existing on the drive
now, use FDISK to create new partitions, and then use FORMAT to format the partitions. This
overwrites any previously existing data on the drive.
5. If you want to recover the data on the drive and no partitions are being shown, you must use a
data recovery program, such as the Norton Utilities by Symantec or Lost and Found by

Power Quest, to recover the data.
6. If all the partitions appear in DISKPART and one is defined as active, run the FIXBOOT command as follows to restore the system files to the hard disk:
FIXBOOT
General File System Troubleshooting for Windows 2000/XP **1355**
7. Type **EXIT** to restart your system. Remove the disk from drive A: or the Windows 2000 CD-ROM from the CD-ROM drive.
8. If you still have the same error as before after you restart your computer, your drive might be improperly configured or damaged.
9. Restart the Recovery Console and run CHKDSK to check for problems with the hard disk.

Troubleshooting Windows 98

Windows 98 is a major improvement in the way computers operate and the way we interface with them. But to expect this operating system to solve all problems and achieve full compatibility with all hardware is not realistic—there are simply too many different hardware and software manufacturers, each with its unique approach, for one operating system to manage. As computer technicians, we must understand both the strengths and weaknesses of computer systems (hardware and software) in order to achieve the best possible performance for our clients. Fortunately, Microsoft has incorporated many tools to help us fine-tune our systems and achieve the best performance. This lesson focuses on some of the methods and tools we can use to identify problems and manage our systems.

How to Fix PRINTER Problem

Windows 95 / 98 printer trouble shooting.

Cause:

A printer not working in Windows 95 / 98 can be caused by any of the following.

1. Printer cable is not connected properly or interrupted.
2. Printer drivers are not installed or are corrupt.
3. Ports are not installed
4. Printer is defective or malfunctioning.

Solution:

Printer cable is not connected properly:

Verify the printer cable is securely connected from the back of the computer to the printer. If you have a device such as a parallel scanner or zip drive temporarily disconnect this drive to verify it is not causing the issue with the printer.

Printer drivers are not installed or are corrupt:

If your printer has recently begin not to work attempt to reinstall the printer drivers to ensure that the printer drivers have not became corrupted or overwritten.

If you are trying to install the printer, verify that you are using the proper installation diskettes and or CD and that Windows is initializing the correct printer.

Ports are not installed:

Verify that the ports are correctly setup within Windows 9x Device Manager. You can check this by opening device manager, clicking the + next to ports and insuring that a Parallel port or LPT1 port is installed.

If this port is not installed, open control panel double click Add new hardware click Next, click No (to not detect new hardware if you have 95) or allow Windows 98 to attempt to detect. If Windows 98 does not detect the new hardware click No and click Next, select Ports and then select to install the LPT / Printer port. This port is generally always LPT1.

Printer is defective or malfunctioning:

If you have read through the general printer trouble shooting section and the above recommendations do not resolve your printer issue it is recommended that you contact the manufacturer of your printer to have the printer replaced or repair.

How to Perform Clean-Boot Troubleshooting for Windows 98

IMPORTANT: This article describes troubleshooting steps that may help you solve problems starting Windows 98. This information is also available in our Windows 98 Startup and Shutdown Troubleshooting Wizard. We recommend using this wizard, but we have also created this text-based article for your convenience.

NOTE: To more effectively search the Microsoft Knowledge Base, use keywords that relate to your issue. If you are searching for troubleshooting information that is not mentioned in this article, search the Microsoft Knowledge Base again by using keywords that are listed in the following Microsoft Knowledge Base article:

Solution:

Start Windows 98 in Safe Mode

If Windows 98 does not start normally, try to start it in Safe mode. Starting Windows 98 in Safe mode bypasses the current real-mode configuration and loads a minimal protected-mode configuration, disabling Windows 98 device drivers and using the standard VGA display adapter. To start Windows 98 in Safe mode, restart your computer, press and hold down the CTRL key until the **Windows 98 Startup** menu appears, and then choose **Safe Mode**. If you can start in Safe mode, see the steps in the "Windows 98 Starts in Safe Mode" section of this article.

If Windows 98 Does Not Start in Safe Mode

Causes & Solution:

- Your computer is infected with a virus. Start computer with clean bootable CD and scan your system with Norton Anti Virus.
- Your computer's CMOS settings are not correct. Check your computer's CMOS settings to make sure they are correct. Note that you may need to contact the computer manufacturer to verify these settings.
- There is a hardware conflict. These conflicts can include, but are not limited to, PCI BIOS settings, IRQ conflicts, redundant COM ports (for example, two COM1 ports, or an internal modem set to use the same COM port as an existing serial port), and defective RAM chips.
- A setting in the Msdos.sys file needs to be changed (for example, the Logo setting should be set to zero).
- You have a compressed drive that is unable to mount a compressed volume file (CVF).
- If you are still unable to start Windows 98 in Safe mode, run the Windows Registry Checker (Scanreg.exe) tool as there may be a problem with the system registry. To start Windows Registry Checker, restart your computer, press and hold CTRL, choose **Command Prompt Only**, type **scanreg**, and then press ENTER.
- If you are still unable to start Windows 98 in Safe mode after you use Windows Registry Checker, install Windows 98 into a new, empty folder. This step helps to establish whether the problem is related to a remnant of the previous operating system (such as a configuration setting) or a hardware problem.

Windows 98 Starts in Safe Mode

If Windows 98 starts in Safe mode, step through the startup process to see if any devices do not load properly:

1. Click **Start**, click **Run**, type **msconfig** in the **Open** box, and then click **OK**.
2. Click **Selective Startup**.
    1. Try different boot options. The following table lists several startup options. The options are labeled Boot A, Boot B, Boot C. To use a boot option, click the appropriate check boxes to select or clear them. Follow the instructions below the table to determine the cause of your problem.

		Boot A	Boot B	Boot C	
4.	--				
5.	Process Config.sys file		Yes	No	Yes
6.	Process Autoexec.bat file		Yes	No	Yes
7.	Process Winstart.bat (if available) Yes		Yes	No	
8.	Process System.ini file	No	Yes	Yes	
9.	Process Win.ini file	No	Yes	Yes	
10.	Load Startup Group items		Yes	Yes	No
11.					

(Row 3 of the printed list holds the column heads: "3. Boot A Boot B Boot C".)

NOTE: The Msconfig tool cannot disable a file that has the read-only attribute, although it behaves as though it can. To determine if the Msconfig tool has replaced the file you are attempting to disable with a copy of the file, text similar to the following text should appear at the beginning of the file:

```
rem
rem   *** DO NOT EDIT THIS FILE! ***
rem
rem   This file was created by the System Configuration Utility as
rem   a placeholder for your AUTOEXEC.BAT file. Your actual
rem   AUTOEXEC.BAT file has been saved under the name AUTOEXEC.TSH.
rem
```

NOTE: If Windows 98 does not start normally under any of the following scenarios, see the "Troubleshooting Protected-Mode Driver Problems" section later in this article.

First, try the Boot A option. If Windows 98 does not start normally under these conditions, try the Boot B option. If Windows 98 does start normally using the Boot A option, there is a problem in the System.ini or Win.ini file. To find which line in the System.ini or Win.ini file is causing the problem:

a. Click S**tart**, click **Run**, type **msconfig** in the **Open** box type, and then click **OK**.
b. Click the **Win.ini** tab.
c. Double-click the Windows folder.
d. Click to clear the **load=** and **run=** check boxes.
e. Click **OK**.
f. When you are prompted to restart your computer, click **Yes**.

If Windows 98 starts normally using the Boot B option, there is a problem with a driver or terminate-and-stay-resident (TSR) program being loaded from the Config.sys or Autoexec.bat file.

If Windows 98 does not start normally with the Boot A or Boot B options, try the Boot C option. If Windows 98 starts normally using the Boot C option, there is a problem with a program that is run during startup.

If you are still unable to start Windows 98 normally, use the System File Checker tool to check for damaged or replaced system files. To start System File Checker, click **Start**, point to **Programs**, point to **Accessories**, point to **System Tools**, click **System Information**, and then click **System File Checker** on the **Tools** menu.

Troubleshooting StartUp Folder Problems:

The problem may be a result of a program that is run during startup. To determine which program is causing the problem:

1. Click **Start**, click **Run**, type **msconfig** in the **Open** box, and then click **OK**.
2. Click the **Startup** tab, and then click each check box to clear it.
3. Click **OK**, and then restart your computer when you are prompted to do so.

If the problem is resolved, follow these steps:

1. Click **Start**, click **Run**, type **msconfig** in the **Open** box, and then click **OK**.
2. On the **Startup** tab, click the first check box in the list to select it.
3. Click **OK**. and then restart your computer when you are prompted to do so.

If the problem is still resolved, repeat steps 1-3, but click the next check box in the list to select it. When the problem returns, the last check box you selected is loading a program that is preventing Windows 98 to start normally. Contact the program's manufacturer for further assistance.

There may also be a problem with a TSR being loaded in the Winstart.bat file (if the Winstart.bat file exists). If the **Process Winstart.bat File** check box is available on the **General** tab in System Configuration Utility, click the check box to clear it, click **OK**, and then restart your computer.

The Winstart.bat file is usually located in the Windows folder, and is used to load TSRs that are required only by Windows-based programs.

Troubleshooting TSR Problems

The problem may be a driver or TSR being loaded from the Config.sys or Autoexec.bat file. To determine if this is the case:

1. Click **Start**, click **Run**, type **msconfig** in the **Open** box, and then click **OK**.
2. Click **Selective Startup**, and then click the **Process Autoexec.bat File** check box to clear it.
3. Click **OK**, and then restart your computer when you are prompted to do so.

If the problem is resolved, the problem driver or TSR is being loaded from the Autoexec.bat file. If the problem is not resolved, the problem driver or TSR is being loaded from the Config.sys file. To determine which line in the Autoexec.bat or Config.sys file is loading the driver or TSR, follow these steps:

1. Click **Start**, click **Run**, type **msconfig** in the **Open** box, and then click **OK**.
2. Click the **Autoexec.bat** or **Config.sys** tab, and then click the check boxes for all non-essential drivers and programs to clear them.
3. Click **OK**, and then restart your computer when you are prompted to do so.

If the problem is resolved, follow these steps:

1. Click **Start**, click **Run**, type **msconfig** in the **Open** box, and then click **OK**.
2. On the **Autoexec.bat** or **Config.sys** tab, click the first check box in the list to select it.
3. Click **OK**, and then restart your computer when you are prompted to do so.

If the problem is still resolved, repeat steps 1-3, but click the next check box in the list to select it. When the problem returns, the last check box you selected is loading the driver or TSR that is causing the problem. Contact the manufacturer of the program for further assistance.

If the problem is not resolved, run the Windows Registry Checker as there may be a problem with the system registry. To start Windows Registry Checker, click **Start**, point to **Programs**, point to **Accessories**, point to **System Tools**, click **System Information**, and then click **Registry Checker** on the **Tools** menu.

Troubleshooting Protected-Mode Driver Problems

The problem may be a Windows 98 protected-mode driver. To determine if this is the case, follow these steps:

1. Click **Start**, click **Run**, type **msconfig** in the **Open** box, and then click **OK**.
2. On the **General** tab, click **Advanced**.
3. Under **Settings**, click a check box to select it.
4. Click **OK**, click **OK** again, and then restart your computer.

If the problem is not resolved, repeat steps 1-4, but click a different check box to select it in step 3. When the problem is resolved, the last check box you selected is causing the problem.

If the problem is not resolved, disable PCI bus IRQ steering in Windows.

If the problem is not resolved, follow these steps to disable devices in Device Manager:

1. Click **Start**, point to **Settings**, and then click **Control Panel**.
2. Double-click **System**.
3. On the **Device Manager** tab, disable all devices under the following branches:

- Display adapters
- Floppy disk controllers
- Hard disk controllers
- Keyboard
- Mouse
- Network adapters
- Ports
- PCMCIA socket
- SCSI controllers
- Sound, video, and game controllers

To disable a device in Device Manager:

- Double-click the branch that contains the device you want to disable, click the device, and then click **Properties**.
- On the **General** tab, click the **Disable In This Hardware Profile** check box to select it, and then click **OK**.
- Restart your computer.

If the problem is resolved, enable the devices you disabled in step 3, and then verify that no devices are conflicting.

NOTE: Enable devices in the following order:

- COM ports
- Hard disk controllers
- Floppy disk controllers
- Other devices

To enable a device and check for possible conflicts, follow these steps:

- Double-click the branch that contains the device you want to enable, click the device, and then click Properties**.**
- On the **General** tab, click the **Disable In This Hardware Profile** check box to clear it.
- Click the **Resources** tab and verify that there are no conflicts listed under **Conflicting Device List**. Note that the **Resources** tab does not appear for each device.
- Click **OK**, and then restart your computer.

If the problem is not resolved, run the Automatic Skip Driver Agent tool to enable any device that has been disabled. To start Automatic Skip Driver Agent, click **Start**, point to **Programs**, point to **Accessories**, point to **System Tools**, click **System Information**, and then click **Automatic Skip Driver Agent** on the **Tools** menu.

If the problem is not resolved, check for a damaged static virtual device driver (VxD) by following these steps:

1. Restart your computer, press and hold down the CTRL key until the **Windows 98 Startup** menu appears, and then choose **Step-By-Step Confirmation**.
2. Press Y at each prompt up to and including the "Load all Windows drivers?" prompt, and then press N to everything else. Note that you should make a list of all the items trying to load after this point. This prevents VxDs from loading and VxDs in the Windows\System\Vmm32 folder from overriding Windows internal VxDs (VxDs built into the Vmm32.vxd file).

If Windows 98 does not start properly, the Bootlog.txt file lists the last driver that loaded successfully, and lists a "LoadFail" entry for each driver that failed to load before the problem occurred.

Installing New Hardware in Windows

This article describes how to install new hardware for use in Windows. This article addresses the following topics:

- **Manufacturer-supplied drivers**
- **Installing modems**
- **Installing printers**
- **Installing a CD-ROM or a Hard Disk**

Manufacturer-Supplied Drivers

To install a manufacturer-supplied driver, follow these steps:

1. In **Control Panel**, double-click **Add New Hardware**.
2. In Windows 95, click **Next**, click **No**, and then click **Next**.

In Windows 98, click **Next**, and then click **Next** again to search for plug and play devices. If the device you are installing is not found, click "**No, the device isn't in the list,**" click **Next**, click "**No, I want to select the hardware from a list,**" and then click **Next**.

NOTE: If Windows 98 finds your hardware when it searches for plug and play devices, click "**Yes, the device is in the list,**" click the device in the list, click **Next**, and then click **Finish**. You do not have to provide a manufacturer-supplied driver.

3. Click type of hardware, which you are installing the driver, and then click **Next**.
4. Click **Have Disk**.
5. Type the path for the driver you are installing and click **OK**, or click **Browse** and locate the driver. You must type the path for or locate the Oemsetup.inf file from the manufacturer.
6. In the dialog box listing the .inf file, click **OK**. Click **OK** to continue.
7. Click the correct driver and then click **OK**.
8. Click **Finish**.

NOTE: If the new hardware is Plug and Play-compatible, you will need to click **Next** and then click **Finish**.

Installing Modems

To install a new modem, follow these steps:

1. In **Control Panel**, double-click **Modems**.
2. If this is to be the first modem installed in the computer, the Install New Modem Wizard starts automatically. If not, click **Add**.
3. If you want Windows to detect your modem, click **Next**. If not, click the "**Don't detect my modem...**" check box to select it, and then click **Next**.
4. If you chose to have Windows detect your modem, Windows queries the serial ports on your computer looking for a modem. If Windows detects an incorrect modem, click **Change**, and select the appropriate manufacturer and model. Click **Next**, and then continue with step 7.
5. If you chose to select your modem manually, click the appropriate manufacturer and model, and then click **Next**.
6. Click the appropriate communications port, and then click **Next**.
7. Click **Finish**.

Installing Printers

To install a new printer, follow these steps:

1. Click the Start button, point to Settings, and then click Printers.
2. Double-click Add Printer, and then click Next.
3. Click Local Printer or Network Printer as appropriate, and then click Next. If you click Network Printer, you are prompted for the network path for the printer. If you do not know the correct path, click Browse, or check with your network administrator. Click either Yes or No in the "Do you print from MS-DOS-based programs?" area, and then click Next.
4. Click the appropriate manufacturer and model for your printer, and then click Next.
5. If you chose to install a local printer, click the correct port and then click Next.
6. Type a name for the printer (or accept the default name), and then click either Yes or No in the "Do you want your Windows-based programs to use this printer as the default printer?" area. Click Next.
7. To print a test page, click Yes. Click Finish.

How to Install and Remove Programs in Windows

Windows provides a consistent method for easily adding new MS-DOS-based, Windows 3.x-based, and Windows 95/98-based software. In addition, Windows 95/98-compliant programs register themselves in the uninstall database and can be easily uninstalled.

Installing Software with a Setup or Install Program

Some CD-ROM-based programs are Windows 95/98 aware and have an Autorun.inf file that launches the Setup program automatically. The following procedure is for those programs without this feature.

1. In **Control Panel**, double-click **Add/Remove Programs**.
2. On the **Install/Uninstall** tab, click **Install**.
3. Insert the floppy disk or CD-ROM containing the Setup program, and then click **Next**.
4. Verify the Setup program to be run, or click **Browse** to locate the Setup program. When you click **Next** to continue, the Setup program is run.

NOTE: Some MS-DOS-based programs must be set up with Windows in MS-DOS mode. If you cannot install an MS-DOS-based program, restart Windows in MS-DOS mode and try to install the program again. Refer to the program's documentation for additional information.

Installing Single Components (.inf Files)

If the program you are installing has a Windows 95/98-compliant information (.inf) file, follow these steps:

1. In **Control Panel**, double-click **Add/Remove Programs**.
2. Click the **Windows Setup** tab.
3. Click **Have Disk**.
4. Type the path for the new program's .inf file, or click **Browse** and locate the .inf file, and then click **OK**.

NOTE: When you are browsing for the .inf file, you can select only the first .inf file. However, all .inf files in the folder are read. You can type the name of the .inf file you want.

TITLE : Add/Remove Programs Selects Only the First Information File

5. Click the check box for the program element you want to install, and then click **Install**.
6. Click **OK** to finish the installation.

Removing a Windows 95/98-Compliant Program

If the program is Windows 95/98 compliant, follow these steps to uninstall it:

1. In **Control Panel**, double-click **Add/Remove Programs**.
2. On the **Install/Uninstall** tab, click the program you want to remove, and then click **Add/Remove**.

Connecting to the Internet in Windows 98

Windows 95 and Windows 98 have built-in support for connecting to an Internet service provider (ISP) by using either a PPP (Point-to-Point Protocol) or SLIP (Serial Line Internet Protocol) account. SLIP support is available on the Windows 95 and Windows 98 CD-ROM, the Extras for Windows 95 upgrade CD-ROM, and the Internet Jumpstart Kit component of Microsoft Plus! for Windows 95.

To connect to the Internet, you need a PPP or SLIP account with an ISP. You also need the following information from your ISP:

- User name
- Password
- Local access phone number
- Your host and domain name
- DNS server IP address

Authentication technique (whether or not a terminal window is used)

Your ISP may also supply you with an IP address and an IP subnet mask. These optional items are needed only if your ISP gives you a dedicated IP address to use every time that you dial in.

How to Set Up Dial-Up Networking

Before you set up your Internet connection, make sure Dial-Up Networking is installed on your computer. To do this, click **Start**, point to **Programs**, point to **Accessories**, and then see if a folder called Dial-Up Networking is available. If you have a Dial-Up Networking folder, view the "How to Install the TCP/IP Protocol" section in this article.

If you do not have a Dial-Up Networking folder, add Dial-Up Networking. To do this:

1. Click **Start**, point to **Settings**, click **Control Panel**, and then double-click **Add/Remove Programs**.
2. On the **Windows Setup** tab, click **Communications** (do not click the check box), and then click **Details**.
3. Click to select the **Dial-Up Networking** check box, and then click **OK**.
4. On the **Windows Setup** tab, click **OK**.

If you want to use a SLIP account to connect to an ISP in Windows 95, install SLIP support. To do this, use the following steps. **NOTE**: You do not have to perform these steps if you are using Windows 98. SLIP support is installed with Windows 98 by default.

1. Click **Start**, point to **Settings**, click **Control Panel**, and then double-click **Add/Remove Programs**.
2. On the **Windows Setup** tab, click **Have Disk**.
3. Insert the appropriate floppy disk or CD-ROM, point to the Admin \ Apptools \ SLIP folder, and then click **OK**.

How to Install the TCP/IP Protocol

After Dial-Up Networking is installed (or if Dial-Up Networking is already installed), install the Transmission Control Protocol/Internet Protocol (TCP/IP) protocol. To do this:

1. Click **Start**, point to **Settings**, click **Control Panel**, and then double-click **Network**.
2. Click **Add** on the **Configuration** tab, click **Protocol**, click **Add**, click **Microsoft** in the **Manufacturers** box, click **TCP/IP** in the **Network Protocols** box, and then click **OK**.
3. If you want to connect only to the Internet, you need only the following network components:

Client for Microsoft Networks Dial-Up Adapter TCP/IP

To remove a component, click the component, and then click **Remove**.

4. After the necessary components are installed, click **OK**, and then restart your computer when you are prompted to do so.

The following list describes the default TCP/IP settings:

- IP address is set to **Obtain an IP address automatically**
- WINS Configuration is set to **Use DHCP for WINS resolution**
- Installed Gateways: None
- DNS Configuration is set to **Disable DNS**

How to Create and Configure a New Connection for PPP Accounts

1. Click **Start**, point to **Programs**, point to **Accessories**, point to **Communications**, and then click **Dial-Up Networking**.
2. If this is the first time that you are using Dial-Up Networking, the Welcome To Dial-Up Networking Wizard is displayed. If the wizard is not displayed, double-click **Make New Connection**, and then follow the instructions to create a Dial-Up Networking connection.
3. After you create the connection, an icon appears in the Dial-Up Networking folder that you can double-click to connect to the Internet.

If your ISP requires you to manually log on to your PPP account after you have connected, you must enable the necessary support. To do this:

1. Right-click the newly created icon in the **Dial-Up Networking** folder, and then click **Properties**.
2. On the **General** tab, click **Configure**.
3. On the **Options** tab, click to select the **"Bring up terminal window after dialing"** check box, and then click **OK**.
4. On the **General** tab, click **Server Type**.
5. Click **PPP: Windows 95, Windows NT 3.5, Internet** or **PPP: Internet, Windows NT Server, Windows 98** in the **Type Of Dial-Up Server** box, and then click to select the **TCP/IP** check box in the **Allowed Network Protocols** box. When you select the TCP/IP check box, you can connect to the Internet faster after you dial in to your ISP. You do not need to select the other options to connect to the Internet.
6. Click **TCP/IP Settings**.

To use resources on the Internet, you must type your DNS server address. You can also type your IP address if your ISP supplied you with a fixed IP address.

How to Connect to Your PPP Account

To connect to your PPP account:

1. Click **Start**, point to **Programs**, point **Accessories**, point to C**ommunications,** and then click **Dial-Up Networking**.
2. Double-click the **Dial-Up Networking** icon that you created.

3. In the **Connect To** dialog box, type your user name and password, and then click **Connect**.

If you clicked the **Bring up terminal window after dialing** check box is the previous section, a terminal window appears in which you type your logon information. After you type your user name and password, click **Continue**. The order in which you log on and the information that you have to provide may vary, depending on your ISP. Contact your ISP if you have any questions about the logon procedure.

After the **Connected To** dialog box is displayed, you are connected to the Internet.

How to Create and Configure a New Connection for SLIP Accounts

1. Click **Start**, point to **Programs**, point to **Accessories**, point to **Communications**, and then click **Dial-Up Networking**.
2. If this is the first time that you are using Dial-Up Networking, the Welcome To Dial-Up Networking Wizard is displayed. If the wizard is not displayed, double-click **Make New Connection** and follow the instructions to create a Dial-Up Networking connection.
3. After the connection is created, an icon appears in the Dial-Up Networking folder that you can double-click to connect to the Internet.
4. Right-click the newly created icon in the **Dial-Up Networking** folder, and then click **Properties**.
5. On the **General** tab, click **Server Type**.
6. Click **SLIP: UNIX Connection in the Type Of Dial-Up Server**, and then click to select the **TCP/IP** check box in the **Allowed Network Protocols** box. When you select the TCP/IP check box, you can connect to the Internet faster after you dial in to your ISP. You do not need to select the other options to connect to the Internet.
7. Click **TCP/IP Settings**.

To use resources on the Internet you must type your DNS server address. You can also type your IP address if your ISP supplied you with a fixed IP address.

How to Connect to Your SLIP Account

1. Click **Start**, point to **Programs**, point to **Accessories**, point to **Communications**, and then click **Dial-Up Networking**.
2. Double-click the Dial-Up Networking icon that you created in the previous section.

3. In the **Connect To** dialog box, click **Connect**.
4. After the modem negotiation procedure, a terminal window is displayed in which you type your logon information. After you type your user name and password in the terminal window, you should receive a message from your ISP that inform you of your IP address for the session. Most providers use a message such as "Your IP address is" or "SLIP session from ###.###.###.### to ###.###.###.###" where the second set of numbers is your IP address.
   The order in which you log on and the information that you have to provide may vary, depending on your ISP. Contact your ISP if you have any questions about the logon procedure.
5. Note your IP address, and then click **Continue**.
6. Confirm that your IP address in the **SLIP Connection IP Address** dialog box is correct by typing the IP address that noted in step 4, and then click **OK**.
   If you are not sure what your IP address is, contact your ISP and ask how to determine your IP address for your SLIP account.

Advantages of PPP over SLIP

If you have a SLIP account, you may want to investigate the possibility of obtaining a PPP account instead. PPP accounts are faster, offer error correction, may not require you to use the terminal window to log on, and do not require you to manually enter your IP address.

Troubleshooting Tips

If you experience problems when you try to connect to the Internet:

- Verify that your server type is correct. For example, if you have a PPP account, verify that the server type is set to PPP and not SLIP or CSLIP.

If you have a SLIP account, verify that you type the correct IP address when you are prompted to do so during the logon process.

If you are using a SLIP account, you experience problems after you connect to the Internet, and you can successfully ping another computer but you cannot use Winsock programs to connect to Internet resources, change the server type from SLIP to CSLIP (or from CSLIP to SLIP).

To verify that your Internet connection is working properly, use the **Ping** command to ping a known **IP address**. To do this:

1. Click **Start**, point to **Programs**, and then click **MS-DOS Prompt**.
2. Type the following command at an MS-DOS prompt, where *IP address* is a known IP address, and then press ENTER:

ping *IP address*

If you can ping a known IP address successfully, basic TCP/IP connectivity is functioning properly. Next, try to ping a known friendly name instead of an IP address. To do this, type the following command at an MS-DOS prompt, where *friendlyname* is the Uniform Resource Locator (URL) for a Web site, and then press ENTER:

ping *friendlyname***.com**

If this works, your DNS settings are functioning properly.

Windows REGISTRY

REGISTRY BASICS

The two major configuration areas are the INI files and the Registry. As mentioned earlier, the INI files are used to a lesser degree in Win95 than in Windows 3.x. The Windows 9.x Registry is a central depository of configuration information. The registry is a database comprised of two files:

> **SYSTEM.DAT ?** contains information about the hardware configuration of the computer.
>
> **USER.DAT ?** contains environmental settings for each user who logs onto the machine.

There is also an optional CONFIG.POL file that you can create. This file contains security settings about which features of Windows 95 are enabled or disabled for a particular user, in other words - the Policy for the user and or for the computer. CONFIG.POL is edited with the Policy Editor, which is located on the Windows 95/98 installation CDROM in the \ADMIN\APPTOOLS\POLEDIT directory.

The USER.DAT and SYSTEM.DAT files can only be edited using the Registry Editor (REGEDIT.EXE), or by clicking Start, Run, REGEDIT. The registry contains 6 keys:

HKEY_CLASSES_ROOT Contains file extension associations.

HKEY_CURRENT_USER User profile information for current person logged on

HKEY_LOCAL_MACHINE Contains settings for hardware installed on the computer

HKEY_USERS Contains default user profile and current session profile

HKEY_CURRENT_CONFIG Contains current hardware configuration

HKEY_DYN_DATA Contains dynamic (in RAM) settings for Plug-n-Play devices

By clicking on the plus sign next to each key, you can open subkeys, or folders, that contain settings that you may wish to change. To find a particular setting, you can either browse to it, or select Edit, then Find from the REGEDIT menu.

> Every time Windows 9x successfully loads the Registry, it makes a backup copy of USER.DAT to USER.DA0 and SYSTEM.DAT to SYSTEM.DA0. If you suspect registry corruption, you can boot into Safe Mode Command Line (F8 at startup) and rename the two registry files from .DA0 back to .DAT and reboot. The registry will be current as of the last successful boot.

Most often, changes you make using Windows 9x configuration utilities, such as the Properties menu option, Control Panel, or Device Manager, automatically make changes to the registry.

Managing Hardware in Windows 9x

i. **Device Manager** **-** The Device Manager is a graphical view of all the hardware installed in your computer that Windows 9x can detect. You can get to device manager by right clicking My Computer, or by selecting the System icon in Control Panel. You can click the plus sign next to a category and it will ?tree out? and then you can see all the devices in a category. Then, if you double click a device, or select Properties of a selected device, then you can see more detailed information on the device on three tabs: General, Driver, and Resources. If a device has a resource conflict, Windows 9x Device Manager will inform you of device resource conflicts, where Windows 3.x would not.

ii. **Updating a Device?**s Properties or Driver ? The Device tab located in the properties of a device in Device Manager is where you can update or install drivers. You can also assign IRQ and or other Resources for a device from the Resource tab.

Windows 9x Management

Windows 9x has three areas where one can improve system performance:

1. **Disk Management - In Windows 9x, the swap file and disk drive caching configuration are now automatic.**

    - **Virtual Memory ?** The swap file is used to provide ?virtual memory?. This is simply hard drive space that is used for idle programs while other active parts of programs are swapped into main memory (RAM). The larger the swap file, the fewer times the computer needs to perform a hard drive search for the needed data. A swap file is never as good as more RAM, but it does help performance. The default setting is for Windows to manage the swap file size and this is the recommended

configuration. It is not a good idea to disable the swap file, as the system can become unbootable.

- **Disk Caching** - Disk caching in Windows 9x is now handled by the 32 bit program VCACHE, instead of SMARTDRIVE, which was used in earlier versions of Windows.

- **Disk Conversion ?** When users upgrade to Windows98 from Win3.x, or Win95, they are probably still using FAT16, because that is what Win3.x and Win95 used. Included with Windows 98 is the FAT32 Drive Converter utility that will enable users to convert their partitions from FAT16 to FAT32, without destroying any data and perhaps gain some more free hard disk space. The utility is located by selecting Start, Programs, Accessories, System Tools and then Drive Converter.

- **Disk Compression ?** The disk compression utility that is included with Windows 9x, DriveSpace, is loaded automatically during a Windows 9x installation. The Windows 9x version of DriveSpace is a protected mode driver with faster performance.

Resource Monitoring

- **Microsoft System Information ?** The System Information tool is a very comprehensive troubleshooting tool for browsing hardware resources, components, software environment and applications. It is a display only utility. Select Start, Programs, Accessories, System Tools, and then System Information.

- **System Monitor ?** This tool is useful to discover what effect a particular application may have on a system. System categories such as Disk Cache, File System, Kernel, Memory Manager and MS Network Client can be monitored.

Shareware ? There are many free utilities and tools available from the internet from such URL?s as www.download.com or www.tucows.com

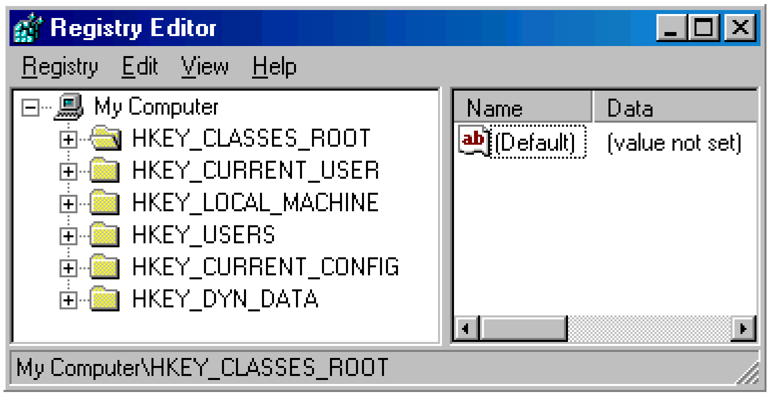
The registry contains extended information, settings and various other values for the the Microsoft Windows 95, Windows 98, Windows NT, Windows 2000, Windows ME, Windows XP Operating System. Within the registry you can control a great majority of the operating system as well as fix a lot of issues with Windows. However we only

recommend those whom are experienced with computers or whom feel confident to go into the registry.

Before going into the Registry and changing or deleting anything we ALWAYS recommend that you **backup the registry**.

The computer registry consists of two files hidden in the Windows directory, system.dat and user.dat. User-specific system information is contained in the user.dat file and Computer and Hardware specific information in the system.dat file.

To get into the Windows 95 or Windows 98 registry click Start / Run / type **regedit**



When typing **regedit** you will get the following screen this window as you can see is the Registry editor. This windows looks very similar to Explorer however within each folder you have sub folders which have various settings.

The following are the types of files that you will see during the time you are in the registry.

- **Open or selected folder** just as you would see in Windows Explorer
- **Closed folder** just as you would see in Windows Explorer
- **String Value** allows you to place certain values to certain aspects of a program such as a version number.
- **Binary Value** allows you to set attributes to a particular application, values are in binary.
- **Dword** Similar to the binary value allowing you to set attributes however done in binary and hex.

BACKING UP / RESTORING THE REGISTRY

Backing up Windows 98 Registry:

To backup Windows 98 Registry from the Windows desktop click Start / Shut Down / Restart the computer in MS-DOS mode.

At C:\WINDOWS> type the following:

attrib **user.dat -r -a -s -h** <press enter>
attrib **system.dat -r -a -s -h** <press enter>
md **backup** <press enter>
copy **user.dat backup** <press enter>
copy **system.dat backup** <press enter>

The above will copy the registry files into a backup directory in your Windows directory. Once the above has been done you can restart the computer and edit the registry as needed.

Restoring Windows 98 Registry:

If you by chance make a mistake while in the registry and have done the above, get to a DOS prompt. If you are not able to get into Windows. As the computer is booting up and you see Starting Windows 98 or you hear a beep press your F8 key on the keyboard this should get you into the Windows 98 Startup menu. Choose the option for Safe Mode command prompt only.

When at the DOS prompt type **cd\windows** (note you will not be able to get into this directory if you have not followed the above steps earlier in backing up your registry.

Once in the Windows directory type the following:

attrib **user.dat -r -a -s -h** <press enter>
attrib **system.dat -r -a -s -h** <press enter>
del **user.dat** <press enter>
del **system.dat** <press enter>

Once these files have been deleted type **cd\windows\backup** once in the windows\backup directory type the following:

copy **user.dat c:\windows** <press enter>
copy **system.dat c:\windows** <press enter>

The above should copy two files, once copied reboot the computer and you should now be able to get back into Windows.

Backing Windows 98 Registry:

Windows 98 has a new feature referred to as scanreg that will automatically backup your registry each time you boot up your computer, however the steps listed above for backing up Windows 98 registry also work but are not necessarily need unless you wish to keep an originally un-touched copy of the registry.

Restoring Windows 98 Registry:

To restore Windows 98 registry get into a DOS prompt, if you are not able to get into Windows 98 to shutdown the computer to get into a DOS prompt as the computer boots up press and hold your left ctrl key on your keyboard which should get you into a Windows 98 startup menu, in this menu choose the option for command prompt only (safe mode command prompt only will not work).

Once at the prompt type **cd\windows\command**

Once in this directory type **scanreg /restore** this will restore a previously backed up copy of your registry. You then should be able to reboot the computer and get back into Windows.

If you backed up your registry using the Windows 98 steps use the Windows 98 restore steps to restore the registry.

REGISTRY KEYS

HKEY_CLASSES_ROOT

Points to branch of Hkey_Local_Machine that describes certain software settings. essential information about OLE and association mappings to support drag-and-drop operations, Windows 98 shortcuts (which are, in fact, OLE links), and core aspects of the Windows 98 user interface.

HKEY_CURRENT_USER

This key points to a branch of Hkey_Users for the user who is currently logged on.

HKEY_LOCAL_MACHINE

Contains computer-specific information about the type of hardware installed, software settings, and other information. This information is used for all users who log on to this computer.

HKEY_USERS

Contains information about all the users who log on to the computer, including both generic and user-specific information. The generic settings are available to all users who log on to the computer. The information is made up of default settings for applications, desktop configurations, and so on. This key contains sub keys for each user that logs on to this computer.

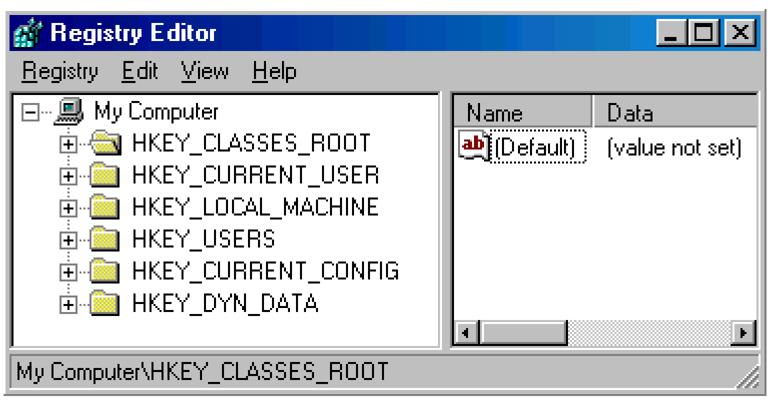
HKEY_CURRENT_CONFIG

Points to a branch of Hkey_Local_Machine\Config that contains information about the current configuration of hardware attached to the computer.

HKEY_DYN_DATA

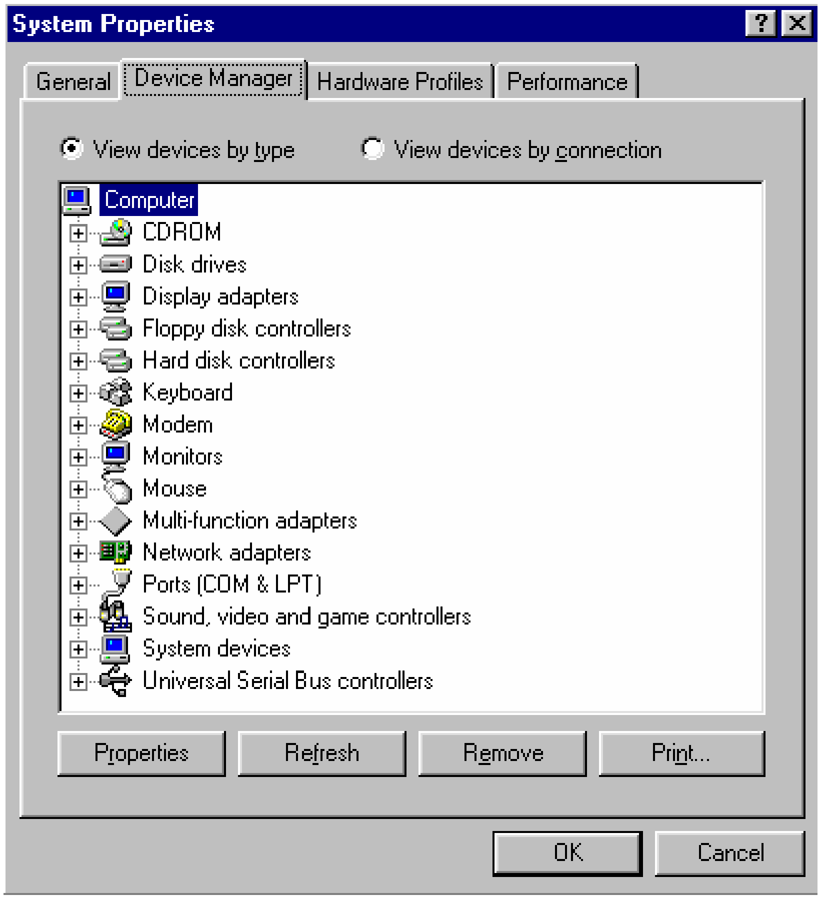
Points to a branch of Hkey_Local_Machine that contains the dynamic status information for various devices as part of the Plug and Play information. This information may change as devices are added to or removed from the computer. The information for each device includes the related hardware key and the device's current status, including problems.

REGISTRY SCREEN SHOTS



The Main Windows of the Registry Editor

WINDOWS DEVICE MANAGER



DEVICE MANAGER ABCs

First introduced in Microsoft Windows 95. Microsoft Windows device manager is a program that allows users to view hardware devices detected in Windows and manage those hardware devices easily. Device manager is available in Windows 95, Windows 98, Windows ME, Windows 2000 and Windows XP.

DEVICE MANAGER ERRORS

Within device manager the user can locate three types of problems an exclamation mark, a red X, or an Other device.

Exclamation Mark - When a device has an explanation mark this is an indication that the device is conflicting with another hardware device.

Red X - Indication that the device has been disabled, removed or that Windows is unable to locate the device.

Other Devices - Windows was unable to locate the drivers for the device and cannot determine the type of device being installed.

DEVICE MANAGER OVERVIEW

To get into the device manager click Start / Settings / Control Panel / double click the System icon and within System Properties click the Device Manager tab or the Hardware tab depending upon which version of Windows you are running.

Users can also open the System Properties by right clicking on the My computer icon on the desktop and click properties.

Finally, you can also access the System Properties if you have a Windows Keyboard by pressing the windows key on the left hand side of the keyboard and the pause key at the same time.

Viruses

Viruses are nasty little programs that can wreak havoc on a computer and its data. The sole purpose of a virus is to replicate itself and make life miserable for computer users. Many viruses are simple annoyances, but some of them can cause irreparable harm to files.

Viruses can be caught from various sources including shareware, files downloaded from the Internet, software from unknown origins, and bulletin boards.

There are four basic types of viruses:

- **File Infectors:** These attach themselves to executable files and spread to other files when the program is run.
- **Boot Sector:** These replace the master boot record (or boot sector on a floppy disk). They write themselves into memory any time the computer is booted.
- **Trojan Horses:** These are disguised as legitimate programs, but when loaded, they begin to harm the system.
- **Macro Viruses:** These common nuisances attach themselves as executable code to documents (such as Microsoft Word documents) and run when the document is opened. (They can also attach themselves to certain kinds of e-mail.) It used to be true that you couldn't get a virus from opening a document; running a program was required. Unfortunately, this has changed thanks to the widespread use of macros by computer users. While macros are very valuable, they mean that when you open a document you *are* running a program.

Viruses have become a way of life in the computer world. With this in mind, there are several measures you can take to prevent, or at least minimize, the damage:

- Purchase a good antivirus program—there are several available. Make sure your choice is compatible with Windows 98. Old MS-DOS antivirus programs do not work well with Windows 98 and might do more damage than good.
- If the computer has a BIOS setting that allows you to disable boot-sector writes (prevent applications from writing to the boot section of the hard disk), enable it! This setting must be disabled before installing Windows 98.
- Viruses are often transmitted by floppy disks. Be careful when reading a floppy disk of unknown origin or using your disk on an unfamiliar machine.
- Currently, many viruses and macroviruses are transmitted over the Internet. Use *extreme* caution when you download files, especially if they come from sources other than a manufacturer's Web site. The most secure protection against Internet-distributed viruses is to make sure you have an antivirus program running at all times (or at least when you're downloading and first running new files).
- Trust no one when it comes to loading programs on your machine. Be aware that any program you load on your computer could contain a virus.
- Keep your antivirus program updated. Hundreds of new viruses are written and transmitted each month.

Chapter 18

Connectivity & Portability

OBJECTIVES

Networks

Portable Computers

Networks

The ability to expand beyond the limit of a single computer in a single office has expanded the reach of the PC to global proportions. Two technologies have driven this expansion: computer networking and the portable computer. In this chapter, we first take a look at how the networks that link up computers on a global scale are put together. Then we examine the portable computer, whose introduction has allowed users instantaneous access to the computing and networking power of all the latest computer technology anywhere they go.

A *network* is defined as two or more computers linked together for the purpose of communicating and sharing information and other resources. Most networks are constructed around a cable connection that links the computers. This connection permits the computers to talk (and listen) through a wire.

Basic Requirements of a Network

In order for a network to function, three basic requirements must be met: it must provide connections, communications, and services.

Connections

Connections include the hardware (physical components) required to hook up a computer to the network. Two terms are important to network connections:

- **The network medium:** The network hardware that physically connects one computer to another. This is the cable between the computers.
- **The network interface:** The hardware that attaches a computer to the network medium and acts as an interpreter between the computer and the network. Attaching a computer to a network requires an add-in board known as a network interface card (NIC).

Communications

Communications establish the rules concerning how computers talk and understand each other. Because computers often run different software, in order to communicate with each other they must speak a "shared language." Without shared communications, computers cannot exchange information, and remain isolated.

Services

A service defines those things a computer shares with the rest of the network. For example, a computer can share a printer or specific directories or files. Unless computers on the network are capable of sharing resources, they remain isolated, even though physically connected.

Networking

Next we look at how the basic elements of connections, communications, and services work together to make networks function properly:

- The connections must operate so that any computer can send or receive electrical signals (data) across the physical media that link them.
- Communications must function so that when one computer sends a message, the receiving computer can listen and understand the message.
- Computers on a network must either provide a service to other computers or make use of a service provided by other computers.

Local Area Networks

A *LAN (local area network)* is a network that covers a limited distance (usually a single site or facility) and allows sharing of information and resources. A LAN can be as simple as two connected computers, or as complicated as a large site. This type of network is very popular because it allows individual computers to provide processing power and utilize their own memory, while programs and data can be stored on any computer in the network. Some of the older LANs also include configurations that rely totally on the power of a mini or mainframe computer (a server) to do all the work. In this case, the workstations are no more than "dumb" terminals (a keyboard and a monitor). With the increased power of today's personal computer, these types of networks are rare.

The primary benefit of a LAN is its ability to share. The following table lists some of the benefits of sharing the most common resources on a LAN.

Resource	Benefit
Data	The sharing of data files that reside in a common location makes multiple-user access easier. Also, it's much easier to maintain data integrity when there is a single, central database. Large customer databases and accounting data are ideal for a LAN system.
Peripherals	Sharing printers, for example, allows more than one user to send jobs to a single printer. This is useful when there is only one high-quality printer in an office and the entire office needs to use it. It also allows

	one user to access multiple printers, providing cost savings in hardware and redundant resources in case one device fails. Other low usage peripherals, such as scanners and plotters, will be better utilized.
Software	Sharing a single copy of an application can be cost-effective (many software manufacturers provide site licenses—licenses for multiple users on a server). It also allows easier maintenance and upgrading.
Storage	Larger, faster disk systems can be used cost-effectively for easy backups.

In addition to the ability to share resources, LANs offer many other benefits that include:

- **Resilience:** Regular backups of the entire system greatly reduce the risk of data loss. Copying data to backup servers allows network operations to continue in the event of primary server failure.
- **Communication gateways:** Low-cost access to fax and Internet connections.
- **Electronic mail:** Cost-effective and convenient communication throughout the network.

Wide Area Networks

A *wide area network (WAN)* spans relatively large geographical areas. Connections for these sites require the use of ordinary telephone lines, T1 lines, ISDN (Integrated Services Digital Network) lines, radio waves, or satellite links. WANs can be accessed through dial-up connections, using a modem, or leased line direct connection. The leased-line method is more expensive but can be cost-effective for transmission of large volumes of data.

Types of Networks

There are essentially two types of networks. They differ in how information is stored, how security is handled, and how the computers on the network interact.

In a *peer-to-peer network*, each computer acts as either a server (sharing its data or services with other computers), or a client (using data or services on another computer) depending on the user's needs. Each user, or workstation, establishes its own security and determines which resources are available to other users. Typically these networks are limited in size (15 to 20 workstations). Microsoft Windows for Workgroups, Windows 95 and Windows 98, Windows NT Workstation, Windows 2000, Novell's NetWare, UNIX, and Linux are some software packages available for peer-to-peer networking.

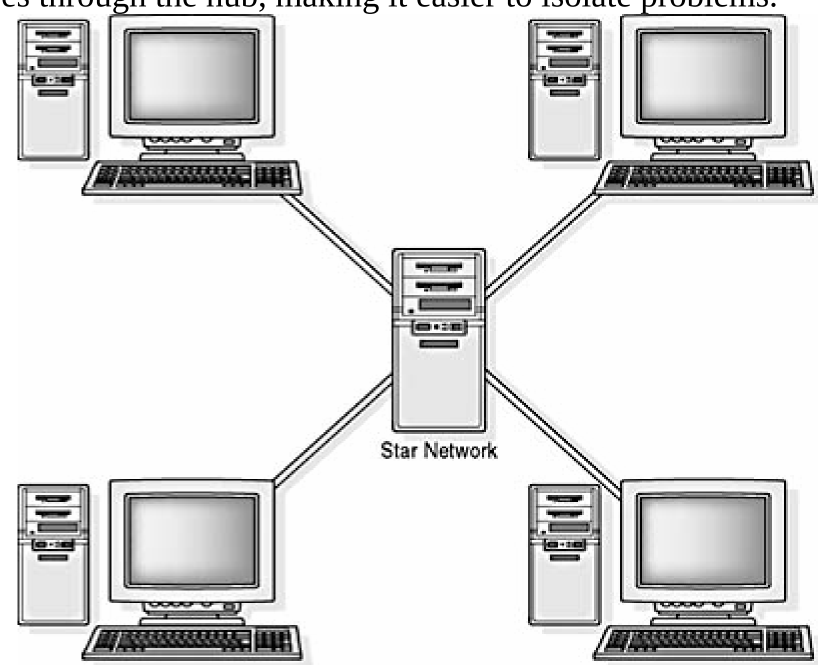
A *server network* requires a central server (dedicated computer) to manage access to all shared files and peripherals. This is a secure environment suitable for most organizations. In this case, the *server* is a computer that runs the network operating system, manages security, and administers access to resources. The *client* is a computer that connects to the network and uses the available resources. The two most common server operating systems are Microsoft's NT4 Server and Novell's IntranetWare. Prior to the release of Windows NT, most dedicated servers worked only as hosts. Windows NT allows the server to also operate as an individual workstation. More than one server can provide services on the network, but only one can be responsible for the security and overall operation of the network.

Network Topology

LAN design is called topology. *Topology* describes the appearance or layout of a network and how data flows through the network. There are three basic types of topologies: star, bus, and ring. In the real world, you are likely to encounter some hybrid combinations of these topologies, but for the A+ exam, we focus only on these three.

Star Topology

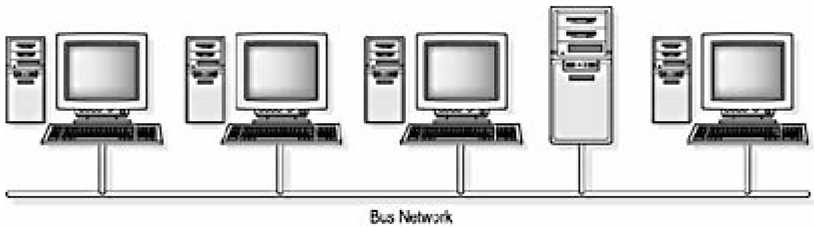
In a *star* network, all devices are connected to a central point called a *hub*. These hubs collect and distribute the flow of data within the network. Signals from the sending computer go to the hub and are then transmitted to all computers on the network. Large networks can feature several hubs. A star network is easy to troubleshoot because all information goes through the hub, making it easier to isolate problems.



Star topology

Bus Topology

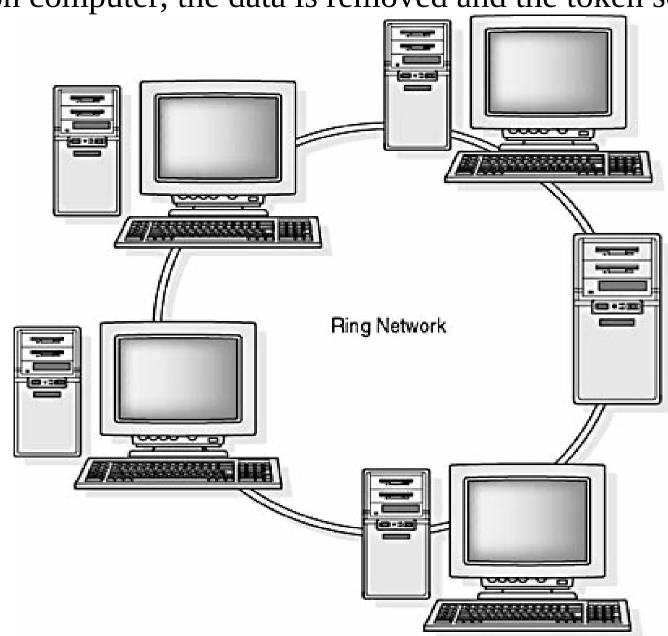
In a *bus* network, all devices are connected to a single linear cable called a trunk (also known as a backbone or segment). Both ends of the cable must be terminated (like a SCSI bus) to stop the signal from bouncing. Because a bus network does not have a central point, it is more difficult to troubleshoot than a star network. A break or problem at any point along the bus can cause the entire network to go down.



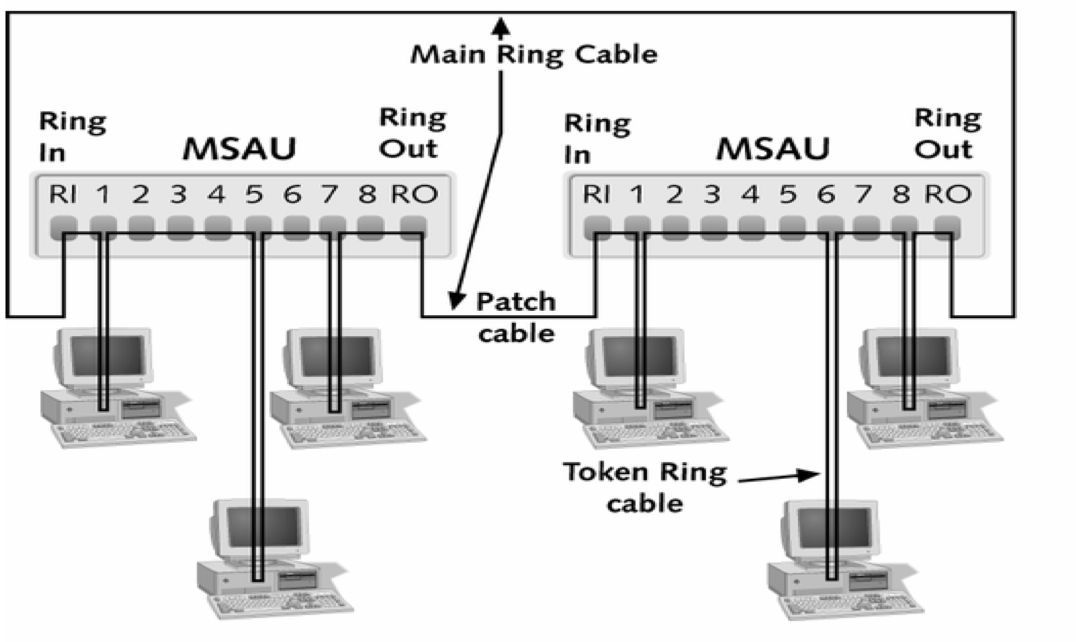
Bus topology

Ring Topology

In a *ring* network, all workstations and servers are connected in a closed loop. There are no terminating ends; therefore, if one computer fails, the entire network will go down. Each computer in the network acts like a repeater and boosts the signal before sending it to the next station. This type of network transmits data by passing a "token" around the network. If the token is free of data, a computer waiting to send data grabs it, attaches the data and the electronic address to the token, and sends it on its way. When the token reaches its destination computer, the data is removed and the token sent on.



Ring topology



Network Operating System (NOS)

The network operating system (NOS) consists of a family of programs that run in networked computers. Some programs provide the ability to share files, printers, and other devices across the network. As previously mentioned, computers that share their resources are called servers; computers that use the resources on other computers are called clients. It is common to run client and server software on the same computer. This enables one user to access the resources on another computer while coworkers make use of resources on your computer.

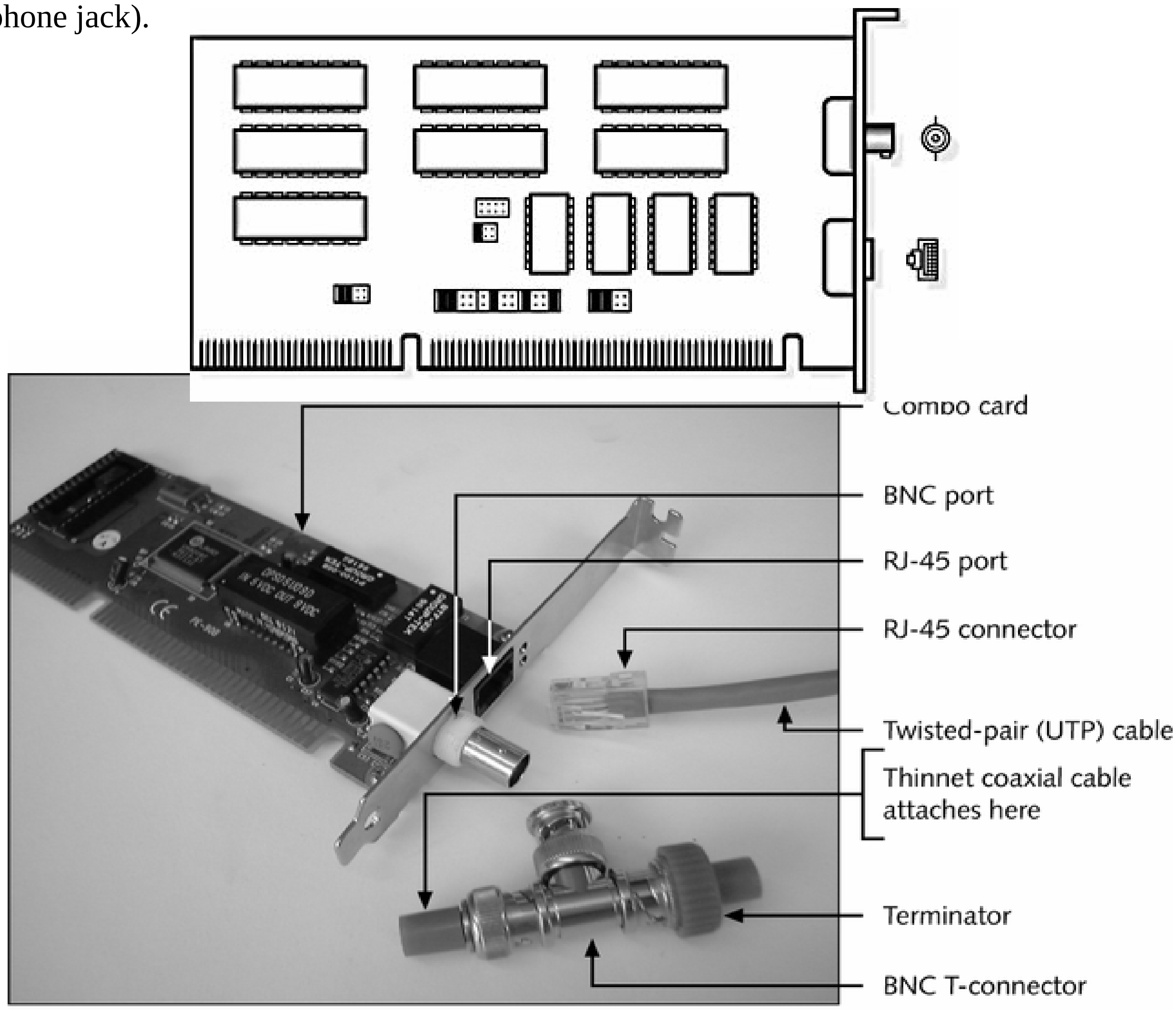
Networking software can be a special program added on to the computer, such as Artisoft's LANtastic or Novell's NetWare, or it can be an integral part of an operating system such as Microsoft's Windows 95 or Windows 98, Windows NT, or Windows 2000.

Network Interface Cards

Network interface cards (NICs) link a computer to the network cable system. They provide the physical connection between the computer's expansion bus and the network cabling. The low-powered digital signals that transmit data inside a computer are not powerful enough to travel long distances. A network interface card boosts these signals so they can cross a network cable. The interface card also must change the form of data from a wide parallel stream—coming in 8, 16, or 32 bits at a time—to a narrow stream, moving 1 bit at a time in and out of the network port (parallel to serial conversion—see Chapter 2.

The network interface card takes data from the computer, packages the data for transmission, and acts as a gatekeeper to control access to the shared network cable. Because the NIC functions as an interface between the computer and the network cabling, it must serve two masters. Inside the computer, it moves data to and from RAM. Outside the computer, it controls the flow of data in and out of the network cable system. Because the computer is typically much faster than the network, the interface card must buffer the data between the computer and cable. This means it must temporarily store the data coming from the computer until it can place it on the network.

Installation of the network interface card is the same as for any other expansion card. It requires setup of the system resources: IRQ, address, and software. Most cards today allow connection for either thin Ethernet or UTP (unshielded twisted-pair) cabling. Thin Ethernet uses a round BNC connector, and UTP uses a RJ-45 connector (similar to a telephone jack).



Network interface card

Installing a NIC is just like installing any other expansion card. If you are installing a Windows 95-compliant Plug and Play card in a Windows 95 or Windows 98 machine, you'll simply need to physically install the card and boot up the computer. The card will be detected and, more than likely, install itself. You might only need to answer a few questions along the way. It requires a little more work to install a NIC in an operating system that is not Plug and Play-compliant. Installing network cards includes the following steps:

1. Be sure to document any changes that you make to the existing computer. This will eliminate any confusion in the installation process and provide future reference in case of problems.
2. Determine whether the card needs IRQ, DMA (direct memory access), or address settings. Remember that you might have to configure these manually, so be sure to check the card's documentation for default settings and instructions for how to make any needed changes.
3. Determine whether the necessary settings are available on the machine on which they will be installed. If proper documentation is not available, use diagnostic software such as Microsoft Diagnostics (MSD) to determine settings. Also check your AUTOEXEC.BAT, CONFIG.SYS, and SYSTEM.INI files; they might give clues as to which settings are already in use.
4. Turn off the machine and remove the cover. Be sure to take all appropriate measures for protection against electrostatic discharge (ESD).
5. Set the NICs jumpers or DIPP (dual inline package) switches as necessary and insert the card.
6. Turn on the machine and run the setup utility provided by the manufacturer. If you are using Windows 95, Windows 98, or Windows 2000, and the NIC is not Plug and Play, you can use the Add New Hardware wizard in the Control Panel to install the drivers and set up the card. (Remember to document all settings.)

Network Cabling

All networks need cables. The three main types are twisted-pair cable (TP), coaxial cable, and fiber-optic cable (FDDI—Fiber Distributed Data Interface).

Twisted-Pair Cable

Twisted-pair cable, shown in the picture consists of two insulated strands of copper wire twisted around each other to form a pair. One or more twisted pairs are used in a twisted-pair cable. The purpose of twisting the wires is to eliminate electrical interference from other wires and outside sources such as motors. By twisting the wires, any electrical noise from the adjacent pair will be canceled. The more twists per linear foot, the greater the effect.

Twisted-pair wiring comes in two types: shielded (STP) and unshielded (UTP). STP has a foil or wire braid wrapped around the individual wires of the pairs; UTP does not. The STP cable uses a woven-copper braided jacket, which is a higher-quality, more protective jacket than UTP.



Twisted pair cable

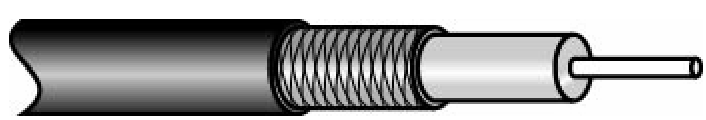
Of the two types, UTP is the most common. UTP cables can be further divided into five categories:

- **Category 1:** Traditional telephone cable. Carries voice but not data.
- **Category 2:** Certified UTP for data transmission of up to 4 Mbps (megabits per second). It has four twisted pairs.
- **Category 3:** Certified UTP for data transmission of up to 10 Mbps. It has four twisted pairs.
- **Category 4:** Certified UTP for data transmissions up to 16 Mbps. It has four twisted pairs.
- **Category 5:** Certified for data transmissions up to 100 Mbps. It has four twisted pairs of copper wire.

Twisted-pair cable has several advantages over other types of cable (coaxial and fiber-optic)—it is readily available, easy to install, and inexpensive. Among its disadvantages are its sensitivity to EMI (electromagnetic interference) and susceptibility to eavesdropping; it does not support communication at distances of greater than 100 feet; and it requires the addition of a hub (a multiple network connection point) if it is to be used with more than two computers.

Coaxial Cable

Coaxial cable is made of two conductors that share the same axis; the center is a copper wire that is insulated by a plastic coating and then wrapped with an outer conductor (usually a wire braid). This outer conductor around the insulation serves as electrical shielding for the signal being carried by the inner conductor. Outside the outer conductor is a tough insulating plastic tube that provides physical and electrical protection. At one time, coaxial cable was the most widely used network cabling. However, with improvements and the lower cost of twisted-pair cables, it has lost its popularity.



Coaxial cable

Coaxial cable is found in two types: thin (*ThinNet*) and thick (*ThickNet*). Of the two, ThinNet is the easiest to use. It is about one-quarter of an inch in diameter, making it flexible and easy to work with (it is similar to the material commonly used for cable TV). ThinNet can carry a signal about 605 feet (185 meters) before the signal strength begins to suffer. ThickNet, on the other hand, is about three-eighths of an inch in diameter. This makes it a better conductor—it can carry a signal about 1,640 feet (500 meters) before signal strength begins to suffer. The disadvantage of ThickNet over ThinNet is that it is more difficult to work with. The ThickNet version is also known as standard Ethernet cable.

When compared to twisted-pair, coaxial cable is the better choice even though it costs more. It is a standard technology that resists rough treatment and EMI. Although more resistant, it is still susceptible to EMI and eavesdropping.

Use coaxial cable if you need:

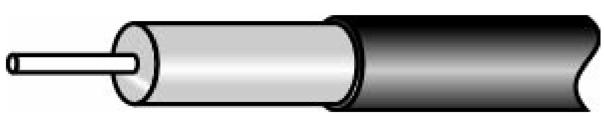
- A medium that can transmit voice, video, and data.
- To transmit data longer distances than less-expensive cabling.
- A familiar technology that offers reasonable data security.

A Mixed-Cable System

Many networks use both twisted-pair and coaxial cable. Twisted-pair cable is used on a per-floor basis to run wires to individual workstations. Coaxial cable is used to wire multiple floors together. Coaxial cable should also be considered for a small network because you can purchase prefabricated cables (with end connectors installed) in various lengths.

Fiber-Optic Cable

Fiber-optic cable is made of light-conducting glass or plastic fibers. It can carry data signals in the form of modulated pulses of light. The plastic-core cables are easier to install, but do not carry signals as far as glass-core cables. Multiple fiber cores can be bundled in the center of the protective tubing.



Fiber-optic cable

When both material and installation costs are taken into account, fiber-optic cable can prove to be no more expensive than twisted-pair or coaxial cable. Fiber has some advantages over copper wire; it is immune to EMI and detection outside the cable and

provides a reliable and secure transmission media. It also supports very high bandwidths (the amount of information the cable can carry), so it can handle thousands of times more data than twisted-pair or coaxial cable.

Cable lengths can run from .25 to 2.0 kilometers depending on the fiber-optic cable and network. If you need to network multiple buildings, this should be the cable of choice. Fiber-optic cable systems require the use of fiber-compatible NICs.

Specifying the Right Cable

In order to ensure trouble-free operation, network cabling must match the system requirements. Cable specifications are based on three factors: speed, bandwidth, and length. Cables are designated with names like 10Base5. Speed is the first number in the identification¾representing the maximum transmission speed (bandwidth) in Mbps. This will be 1, 5, 10, or 100. Band is the second part of the identification. It is either base or broad depending upon whether the cable is baseband or broadband. The last part of the identification refers to the cable length or cable type. If the unit is a number, it is the maximum length of the cable segments in hundreds of meters (1 meter is approximately 3.3 feet). In some cases, it can refer to 50-meter increments (1Base5 is five 50-meter increments—250 meters). In other cases, it represents cable type: T (twisted-pair) or F (fiber-optic). The following table shows the common types of cables and their specifications.

Name	Description	Type	Segment	Speed
10BaseT	Common	UTP twisted-pair	.5 to 100 meters	10 Mbps
10Base2	Ethernet ThinNet	Coaxial	185 meters	10 Mbps
10Base5	Thick Ethernet	Coaxial	500 meters	10 Mbps
100BaseT	Becoming common	Twisted-pair	.5 to 100 meters	100 Mbps

The preceding table covers the basic cable requirements for the A+ networking objective; however, there are many other forms of network connections. For example, you'll find microwave links; forms of radio; and, for small offices and homes, power-line networks (whose NICs have connectors that plug into wall sockets, allowing regular wiring to carry the signal), and telephone-line networks that use standard phone jacks to plug into existing lines. These have relatively short ranges (generally limited to one office or one floor of a building).

LAN Communication

A LAN is similar to a telephone system with one party line—not everyone can talk at the same time. The difference is that, with a LAN, the speed is so fast that it fosters the perception that many transactions are taking place at the same time. But just like a one-lane road, the heavier the traffic, the slower it moves.

Ethernet

Ethernet uses a system known as carrier sense multiple access with collision detection (CSMA/CD). It also uses the bus topology discussed earlier in this lesson. The term "carrier sense" means that the network card listens to the cable for a quiet period during which it can send messages. "Multiple access" refers to the fact that more than one computer can be connected to the same cable. And "collision detection" is the ability to detect whether messages have collided in transit (in which case neither message will arrive at its destination and both will be retransmitted).

Fast Ethernet was developed to meet the increasing demands on networks. Fast Ethernet works on the same principals as the original Ethernet, but operates at 10 times the speed. Ethernet transmits at 10 Mbps, and Fast Ethernet transmits at 100 Mbps.

Token Ring

As described earlier, a token ring network uses a "token" as the basis for deciding who can communicate on the network. Token rings transmit at 4 or 16 Mbps.

Network Protocols

A network protocol is a set of rules that govern the way computers communicate over a network. In order for computers using different software to communicate, they must follow the same set of networking rules and agreements, called *protocols*. A protocol is like a language; unless both computers are speaking and listening in the same language, no communication will take place.

Networking protocols are grouped according to their functions, such as sending and receiving messages from the NIC, or talking to the computer hardware and making it possible for applications to function in a network. Early computer networks had manufacturer-unique inflexible hardware and strict protocols. Today's protocols are designed to be open, which means they are not vendor-, hardware-, or software-specific. Protocols are generically referred to as protocol families or protocol suites because they tend to come in groups (usually originating from specific vendors).

The following is a list of standard network protocols:

- IPX/SPX (Internetwork Packet Exchange/Sequenced Packet Exchange): The NetWare core protocol developed by Novell in the early 1980s.
- NetBIOS/NetBEUI (Networked Basic Input/Output System/NetBIOS Enhanced User Interface): A local area protocol developed by IBM and refined by Microsoft; originally, the native protocol for LAN Manager and Windows NT. IBM developed NetBIOS as a way to permit small groups of computers to share files and printers efficiently. NetBIOS is the original edition; NetBEUI is an enhanced version for more powerful networks based on 32-bit operating systems.
- TCP/IP (Transmission Control Protocol/Internet Protocol): A set of standard protocols and services. It was developed by the Department of Defense beginning in the early 1970s as part of an effort to link government computers. This project led to the development of the Internet. Because TCP/IP is the foundation of the Internet, as well as the most widely used networking protocol, it is a good choice for networks.
- AppleTalk: A networking protocol utilized by Macintosh computers.
- DLC (Data Link Control) protocol: The oldest protocol of this group. IBM developed DLC to connect token-ring-based workstations to IBM mainframe computers. Printer manufacturers have adopted the protocol to connect remote printers to network print servers.

Depending on the operating systems and the function of the network you work on, you will probably use more than one network protocol. It's important to get and install LAN drivers that can switch between one protocol and another as needed. The aforementioned protocol information provides you with a rudimentary understanding of basic network techniques and terminology. However, networks are a very complicated subject, and additional training resources should be obtained before installing a network on your own.

Extending a LAN

The previous section on network cables mentioned some limits to the length of cables. The requirements of today's LANs will often exceed the capability of these cables. The following table lists several devices that can be used to extend a LAN network beyond its normal limits.

Devices	Description
Repeaters	The main purpose of a repeater is to extend the length of a network beyond its normal cable lengths. A repeater works like an amplifier to increase or boost the signal to allow transmissions over longer distances. Repeaters are used to connect network segments (groups of computers on the same network). They can also be used to connect segments composed of different media (for instance, a ThinNet segment

	to a fiber-optic segment).
Bridges	Bridges work like repeaters, but offer additional advantages. They can isolate network traffic or problems. Should any problems occur within one segment, the bridge will isolate that segment and not affect other segments on the network, thereby reducing the load on the network as a whole. Bridges can also link segments that are unalike (such as Ethernet and token ring).
Routers	Routers provide interconnectivity between like and unlike devices on the LAN and WAN. Routers work like bridges, but can connect networks using different protocols. They are able to select the best route from one network to another network based on traffic load. Routers determine the flow of data based on such factors as least-cost, minimum delay, minimum distance, and least congestion. Routers are generally used to create a WAN and connect dissimilar networks.
Gateways	Gateways provide all the connectivity of, and even greater functionality than, routers and bridges. A gateway usually resides on a dedicated computer that acts as a translator between two completely dissimilar systems or applications. Because gateways are both translators and routers, they tend to be slower than bridges or routers. Gateways also provide access to special services such as e-mail or fax functions.

Maintaining and Troubleshooting Networks

Maintaining and troubleshooting networks differ according to the operating system. Therefore, you will need to refer to the operating systems' manuals for detailed troubleshooting procedures. A thorough understanding of network troubleshooting is not a requirement of the A+ Certification program. (The section that follows describes some advanced certification programs that focus on networks.) As an A+ technician, you should be familiar with some generic troubleshooting concepts as presented in the following table.

Situation	Probable Cause
Reduced bandwidth	Called a bottleneck, this occurs when the network doesn't handle as much data as usual. A bottleneck is some constraint that limits the rate at which a task can be completed. If a task uses the processor, network, and disk resources, and spends more of its time transferring data to and from the disk, you

	could have a memory bottleneck. A memory bottleneck might require additional RAM.
Loss of data	If data transfers are incomplete or inaccurate, check to ensure that all network cabling and connectors are intact.
Slow loading of programs and files	Fragmentation (see Chapter 8) occurs when the operating system saves, deletes, and moves information. You must defragment the drive. If slow loading persists even after defragmenting, check for memory bottlenecks.
Unauthorized software	You must manage software distribution to ensure that users are not loading unlicensed software and computer viruses on the network. One way is to load only software from a centralized location or server and then remotely copy it to local hard disk drives.
Traffic overloads	A hardware or software failure can bring a LAN to a halt, or the failure can result in more data traffic than the network is designed to handle. You might receive an error message or you might not see any signs other than poor network performance. You must have a system in place that can monitor and manage network traffic. To resolve this problem, you will need to reduce the traffic on the LAN or expand its capabilities.
Common mode failures	Some LAN-component failures affect other components. This is known as a common mode failure. For example, the on-board logic of a NIC might jumble the data format. The NIC will hand the result to the network operating system, which might not detect the error. If the network operating system puts that data into a file, the file will become corrupt.
Network-security violations	Entire books address the subject of network security alone. Every operating system is different, and every customer requires a different level of security. First determine the customer's needs, and then find and read the appropriate documentation.

Network Certification

This chapter is designed to give you a foundation in networks and a general understanding of network design and applications. Technician certification is a growing trend in the computer industry. The A+ examination touches on network terminology and

design; however, some of the most popular networking certification programs are available through Microsoft and other NOS manufacturers. These companies offer many levels of certification; you should consult check manufacturer Web sites and community colleges for detailed course contents. Let's take a look at some of the available programs.

Microsoft Certified Product Specialist (MCPS)

Microsoft Certified Product Specialist (MCPS) certification is designed for advanced end users, computer service technicians, and network administrators who seek the fastest-growing certification in the computer industry. The MCPS certification is useful for those who want to demonstrate expertise with a particular Microsoft product, such as Windows NT Server or Windows NT Workstation. This is also the first step toward becoming an MCSE.

Microsoft Certified Systems Engineer (MCSE)

The Microsoft Certified Systems Engineer (MCSE) certification is the most sought-after certification in the computer industry. Qualified MCSEs plan, implement, maintain, and support information systems in a wide range of computing environments using the Microsoft Windows NT Server and the Microsoft BackOffice integrated family of server products. To become an MCSE, you must pass four core modules and two elective exams.

Certified Novell Administrator (CNA)

The Certified Novell Administrator (CNA) certification is frequently the first credential earned by NetWare career professionals. CNA training provides you with the critical day-to-day maintenance and management skills you need to survive in the world of Novell NetWare and IntranetWare. The CNA certification is the first step to becoming a Certified Novel Engineer (CNE).

Certified Novell Engineer (CNE)

The Certified Novell Engineer (CNE) certification is currently one of the most popular credentials in the field of networking. It can give a tremendous boost to the career of any serious networking professional. Novell's certification curriculum is 50 percent industry-generic—as a Certified Novell Professional, you are qualified to support Novell-specific products as well as non-Novell products. Novell certifications are recognized worldwide as the standard of excellence for supporting Novell's Internet, Network, and GroupWare products. One of the leading vendor-certification programs in the IT industry, CNE training gives you the skills to provide high-end, solutions-based technical support.

The Internet

The *Internet*, also known as "the Net," is the most WAN in the world—a network of networks working together. This relatively new communication technology has begun to affect our lives as significantly as television and the telephone. When most people talk about using the Internet, they talk about which Web sites they have visited or who they've met online.

Most LANs make use of passwords and other forms of security, but the Internet is one of the most open networks in the world. Some common Internet uses include communication; locating lost friends and family; researching information for school or work; and locating businesses, products, or services (such as travel). The Internet can be your most valuable resource for virtually anything and everything.

A thorough knowledge of the Internet and how it works is not a requirement of A+ Certification. However, as an A+ technician, you might find that it becomes the single most valuable information tool at your disposal.

Internet Basics

The Internet is really a collection of services. Let's take a look at the most important services and the major concepts behind them.

The World Wide Web (WWW)

When people say they were "surfing" the net, they were probably visiting the collection of hyperlinked Web sites known as the *World Wide Web*. These Web sites are located around the world, and their numbers continue to grow by the thousands every day. Each Web site within the World Wide Web has a unique address called a Universal Resource Locator (URL).

NOTE

The World Wide Web (usually abbreviated as "the Web") is not the Internet—it is only part of the Internet. Although it is currently the largest, most popular, and fastest growing part of the Internet, it represents only a fraction of Internet services available that include FTP, Gopher, and Telnet.

Electronic Mail (E-mail)

Electronic mail, usually known as e-mail (sometimes spelled E-mail or email), is the most commonly used function of the Internet, allowing users to send and receive messages (and files) electronically to and from millions of people all over the world.

Electronic mailing lists allow users to join group discussions with people who share their interests. Like regular mail (also called snail mail), e-mail is also sent to an address (a virtual one). like **Mujahid@domainXcess.com** (here Mujahid is the user & DomainXcess.com is the website, @ symbol represents the e-mail address.)

File Transfer Protocol (FTP)

The *File Transfer Protocol (FTP)* is a special application used for uploading and downloading files to and from the Internet.

Transmission Control Protocol/Internet Protocol (TCP/IP)

The *Transmission Control Protocol/Internet Protocol (TCP/IP*) is the language (network protocol) used by computers to talk to each other over the Net. TCP/IP has also become a common protocol for LANs. Regardless of which operating system or software is being used, your commands travel through the Internet in TCP/IP format. The services of the Internet and the World Wide Web could not be provided without TCP/IP.

Internet Protocol (IP) Address

Each machine on a network is given a unique 32-bit address. These addresses are normally expressed in decimal values of 4 bytes, separated with periods; for example, 127.0.0.1. Without a unique address, there would be conflicts and chaos. This is the same concept as the hardware addresses discussed in Chapter 10, except that it is a software address. It designates the location of its assigned device (usually a NIC) on the network.

Internet Service Providers (ISPs)

Internet service providers furnish the connection between dial-up (modem) users and the Internet. While some are big names with millions of users, there are many more that serve local areas with both dial-up and hosting plans.

Uniform Resource Locators (URLs)

As mentioned, the *Uniform Resource Locator* is the Web's address system. To access a Web site, the user must enter the designated URL on the network. Each URL begins with the character sequence "http://". The letters "HTTP" are an acronym for the Hypertext Transfer Protocol, which identifies the Web site as an address. The rest of the URL is the name of the site. For example Microsoft's URL is http://www.DomainXcess.com. (Because it is universal, it is seldom necessary to first type the characters "http://" when typing a URL in a search engine; most engines take it for granted.)

Domain

An Internet *domain* is a site with a common general interest or purpose, often run by a single firm or institution. The domain suffix gives a general idea of the site's purpose: .com, for businesses, or .edu, for educational institutions. The following table lists common Internet domains.

Domain	Description
.com	Commercial organizations
.net	Networks (the backbone of the Internet)
.edu	Educational institutions
.org	Nonprofit organizations
.gov	US Government nonmilitary institutions
.mil	US Government armed services
com.pk	Two-letter country code.

Domain-Name Server

A *domain-name server* is a computer that matches IP addresses with domain names. The domain name makes it possible for you to use the easy-to-remember domain name DomainXcess.com without having to memorize the string of numbers in the IP address.

Getting Connected

These days, many computer professionals spend a lot of their time getting clients online.

The first thing you need to do to get a client connected is to make sure they have a service provider (we're assuming your client has a computer and a modem). Most people connect to the Internet using independent ISPs that provide local community-based service to Internet users, but popular national ISPs such as The Microsoft Network (MSN) are useful if you travel because many of them have 800 numbers for dial-up access, or many local numbers throughout the country. Local ISPs are great for customers who are looking for a cost-effective company that offers local (including technical) support.

You also need to consider which browser(s) to use. Most ISPs (especially the local ones) provide only the connection or gateway to the Internet. Others provide their own browser software package. Most ISPs allow you to use your choice of browsers. Some local and national ISPs provide a startup CD that includes their recommended browser, as well as FTP tools and other Internet utilities. From time to time, Web surfers will encounter pages that work only with a specific browser. In this case it might be necessary to install both browsers.

Portable Computers

Portable computers, once a novelty, are now a part of everyday business life. Portable computers work and act just like the big ones except they are very compact. In this lesson, we look at those elements that make a portable computer unique.

The category of portable computers includes laptop, notebook, and subnotebook (palmtop) computers, as well as the newest categories: PDA (personal digital assistant) and handheld computers.

Types of Portables

Portable computers are classified according to size and function. Today there are three basic types of portable computers: laptops, notebooks, and subnotebooks.

The first "portable" computers were often called "luggables." The size of a portable sewing machine, they tipped the scales at 30 pounds. Equipped with a small CRT display, they were actually a traditional PC in a slightly smaller case. The real change in portable computers came with the advent of the flat-panel display, allowing the portable to take on the now-familiar slim design. "Laptop" is the term used for the heavier version, usually offering most of the features of a full-fledged PC, but with a folding flat-panel display and integrated keyboard. Notebooks are slender devices that often lack the full range of storage as part of the normal configuration. PDAs, a special group of products offering a subset of features including e-mail, schedule-tracking, contact records, and allowing limited note-taking and Web-browsing, are beyond the scope of this chapter.

Laptop Computers

With advancements in battery technology and the advent of functional large-screen LCDs (liquid crystal displays), the first truly portable computers, referred to as *laptops*, were produced in the late 1980s. These units featured integrated AT-compatible computer boards, including I/O and video controller functions. Laptops, as mentioned, usually feature a folding LCD display and a built-in keyboard and pointing device. They also use an external power supply and a removable, rechargeable battery. Today's laptops have fairly large (2 GB or more) hard drives, a CD-ROM drive, and floppy disk drive (often the latter two are interchangeable plug-ins).

When laptops originally appeared on the market, they were the smallest portable computers made. Today, they are high-end machines that offer features and performance comparable to a desktop system.

Notebook Computers

Advancements in integrated circuit (IC) technology allowed the size of computer components to be reduced even further, and in the early- to mid- 1980s the *notebook computer* was born. Notebooks are roughly 8.75 inches deep, by 11 inches wide, by 2.25 inches thick, and designers are working to decrease the size and power consumption of these units even further. The reduction in size comes at a cost, however, and notebooks typically have smaller and less-capable displays and keyboards than laptops. A wide variety of specialty items have appeared on the market intending to overcome some of the notebook's shortcomings. Docking ports are one such item.

Docking Ports

Docking ports (also known as docking stations) are specialized cases into which an entire notebook can be inserted. This allows the notebook to be connected to desktop I/O devices such as full-sized keyboards, CRT monitors, and network connections. At the very minimum, a docking station provides an AC power source for the notebook. Docking stations are highly proprietary items that are designed for use with specific computer models. They are handy for the user who wants to maintain only one computer system and avoid the necessity of transferring information between two systems. With a docking port and a well-equipped notebook computer, it is possible to have the best of both worlds.

It is not necessary to have a docking port to use a portable computer with a full-sized keyboard, pointing device, and monitor. Most portables have standard connectors for these peripherals. Be aware that you might have to connect the devices before booting up the computer, though.

Subnotebook (Palmtop) Computers

Even smaller than the notebook computers are subnotebook computers, also known as palmtops. These tiny systems are 7 inches wide, by 4 inches deep, by 1 inch high. Due to their size, they are rather limited in function. Keyboards, for example, are too small to permit touch typing. With notebooks decreasing in cost and weight, palmtops have been losing market share and popularity.

Computer Cards (PCMCIA)

To provide laptop and notebook computers with the same expandability associated with desktop computers, the Personal Computer Memory Card International Association (PCMCIA) established several standards for credit-card-sized expansion boards that fit into small slots on these smaller machines. PCMCIA is also referred to as the PC Card bus. The PCMCIA standards have revolutionized mobile personal computers, providing them with the ability to add memory expansion cards, SCSI devices, communication

hardware (for instance, modems and faxes) and many other devices that were previously unavailable to laptop and notebook computer users.

Compatibility problems surfaced along with the development of the PCMCIA card for portable computers. To overcome these incompatibilities, PCMCIA standards were created. The following table outlines the four PCMCIA types and their guidelines.

TYPE	Standard Description
Type I	This original computer-card standard is now referred to as the Type I standard. These slots work only with memory expansion cards. Type I cards are 3.3 mm thick.
Type II	Type II cards support most types of expansion devices (like communication hardware) or network adapters. Type II can accommodate cards that are 5 mm thick.
Type III	Type III slots are primarily for computers with removable hard disk drives. This standard was introduced in 1992. They are 10.5 mm thick; however, they are compatible with Type I and Type II cards.
Type IV	Type IV slots are intended to be used with hard disk drives that are thicker than the 10.5 mm Type III slot.

The PC Card itself is usually sealed in a thin metal case. One end contains the interface to the PCMCIA adapter (68 tiny pinholes); the other end might contain a connector for a telephone line, a network, or another external device.

PCMCIA (PC Card) is part of the Plug and Play standard¾which means it allows you to add components without first shutting off or rebooting the computer. In short, PCMCIA buses are not configured with jumper settings (because they don't have any) but with software.

Portable Computer Hardware

Although many components in a portable computer are similar to those of a desktop system, some components are very different. The major difference between a portable system and desktop system is the display screen.

Displays

Portable computers have a flat, LCD screen that is about .5 inch thick. The display is typically the most expensive component in a portable system. Often it is more

economical to replace the entire computer than to replace the screen. An LCD display is designed to operate at a specific resolution because the size of the pixels on an LCD panel cannot be changed. On a desktop system, by contrast, the signal output from the video adapter can change the resolution on the monitor, thereby changing the number of pixels on the screen. An LCD panel should be thought of as a grid ruled to a specific resolution. Transistors control the color that is displayed by each pixel. The two major types of LCD displays used in portable systems today (dual-scan and active-matrix) are defined by their arrangement of transistors.

Dual-Scan Displays

The dual-scan display (also known as a passive matrix display) consists of transistors running down the *x* and *y* axis of the screen. The number of transistors determines the screen's resolution. Each pixel on the screen is controlled by the two transistors that intersect on the *x* and *y* axis.

If a transistor fails, the entire line of pixels is disabled, leaving a black line across the screen. There is no way to repair this problem except to replace the display. The term "dual-scan" is derived from the fact that the processor redraws half of the screen at a time, which speeds up the refresh rate a little.

Dual-scan displays are considered inferior to active-matrix screens because they tend to be dimmer. They work by modifying the properties of reflected light rather than generating their own light. They are also more prone to ghost images, and make it difficult for two people to see the screen at the same time, because these displays can't be viewed well from an angle. The standard size for this type of screen is 10.5 inch (measured diagonally) with a resolution of 640 by 480. New systems are available with 12.1 inch displays that have a resolution of 800 by 600.

Active-Matrix Displays

Active-matrix displays are also known as thin film transistors (TFTs). They differ from dual-scan screens because they have a transistor for every pixel on the screen rather than just at the edges. Voltages are applied by electrodes at the perimeter of the grid to address each pixel individually.

Because each pixel is powered individually, generating its own light and the appropriate color, a much brighter and more vivid picture results. Creating light instead of altering reflection provides a wider viewing angle, which allows more than one viewer to see the screen at a time. The refreshes are faster and lack the fuzziness associated with the dual-scan systems.

Naturally, the cost of having 480,000 transistors instead of merely 1,400 (on an 800 by 600 screen) makes the active-matrix screen more expensive. Another drawback is that it also requires a lot more power and drains batteries faster. Failure of a transistor causes individual "dead pixels," but this is far less noticeable than the black line caused by a transistor failure of the dual-scan screen.

The 12.1-inch screen has become the standard on high-end laptops with resolutions running at 800 by 600, or even 1,024 by 768. Many portable systems today also include PCI bus video adapters. These screens come very close to the quality of a desktop display.

Screen Resolution

An LCD display's resolution is determined as much by the screen hardware as by the drivers and amount of installed video memory. Some portables can use a "virtual screen" to achieve resolutions of 800 by 600 (and even more) on a 640 by 480 pixel screen. The larger display is held in video memory while the actual screen displays the portion that fits into a 640 by 480 window. The cursor can be used to "pan" the image so that the 640 by 480 window is moved around within the 800 by 600 display. Some manufacturers advertise an 800 by 600 display while using this method, which is a little misleading.

Like a desktop system, color depth is affected by video memory. To operate any LCD display in 16-bit or 24-bit color mode, you must have sufficient video memory available. Portables usually have the video adapter hardware permanently installed on the motherboard, which makes an upgrade virtually impossible. A few PC Card video adapters, however, allow you to connect to an external monitor and increase your video capabilities.

NOTE

LCD technology has progressed to the point that large, flat-panel LCD-type displays are now available for desktop computers, although they're quite expensive.

Processors

Computer CPU manufacturers spend a great deal of time and effort on the design and creation of chips specifically for the portable market. In desktop systems, CPU heat is dissipated by cooling fans housed inside the case. There is no room for this solution in a portable system, so manufacturers have addressed this problem in the packaging of the chip itself.

Chip manufacturer, Intel's, solution to the size and heat problems is the *Tape Carrier Package*. This method of packaging reduces the size, power consumption, and heat

generated by the chip. A Pentium mounted on a motherboard using Tape Carrier Packaging is much smaller and lighter than the pin grid array (PGA) used in desktop systems. The 49-millimeter (mm) square of the PGA is reduced to 29 mm, the thickness to approximately 1 mm, and the weight from 55 grams to under 1 gram.

The Tape Carrier Packaging processor is bonded to a piece of polyamide film (which is like photographic film) using *tape automated bonding (TAB)*. This is the same process that is used to attach electrical connections to LCD panels. The film (called tape) is laminated with copper foil etched to form the leads that connect the processor to the motherboard. When the leads are formed, they are gold-plated to protect them against corrosion, bonded to the processor chip itself, and then the entire assembly is coated with a protective resin.

After being tested, the tape is cut to the proper size and the ends folded into a "gull wing" shape that allows the leads to be soldered to the motherboard while the processor is suspended slightly above it. A thermally conductive paste is inserted between the processor chip and the motherboard, allowing heat to be dissipated through a sink on the underside of the motherboard, while keeping it away from the soldered connections. Of course, because Tape Carrier Packaging processors are soldered to the motherboard, they usually cannot be upgraded.

Some manufacturers use standard PGA processors, sometimes accompanied by fans. As well as a greatly reduced battery life, these systems can be too hot to touch comfortably. Always check the exact model of processor that is used in a system you intend to purchase, not just the processing speed. You might not want to purchase a non-Tape Carrier Packaging processor for the aforementioned reasons.

Voltage Reduction

Mobile Pentiums have operated at 3.3 volts from the days of the original 75-MHz chip, but the newer and faster models have reduced the voltage to only 2.9 volts for internal operations, while retaining the 3.3 volt interface with the motherboard. This translates into a processor that uses as little as 60 percent of the power of a desktop system.

Memory

As with desktop systems, adding memory is one of the most common upgrades performed on portable computers. Unlike desktop computers, which offer only three basic types of slots for additional RAM, there are dozens of different memory-chip configurations designed to squeeze memory upgrades into the small cases of the portable systems.

Some portables use memory cartridges that look a lot like PC Cards, but they plug into a dedicated IC memory socket. Others use extender boards like the SIMMs and DIMMs. In any case, it is strongly recommended that you only install memory modules that have been designed for your system, and only in the configurations recommended by the manufacturer. This does not necessarily limit you to products made by your system's manufacturer, however, because a number of companies manufacture upgrade modules for dozens of systems.

Portable computers use the same types of DRAM and SRAM as desktops and, thanks to advances in thermal management, today's high-end portable systems usually include SRAM cache memory.

Hard Disk Drives

Except for their size and packaging, portable hard disk drive technology is mostly similar to desktops. EIDE drives are standard in portable computers with the exception of the Macintosh computer, which uses SCSI. Internal hard drives, depending on the size of the system, are typically 12.5 mm or 19 mm tall, and use 2.5-inch platters. As with memory modules, hard drives are also mounted in the system a little differently by manufacturers. And, as with memory modules, this can cause upgrade compatibility problems.

Some manufacturers use a caddy to hold the drive and make connections to the system. This makes upgradability as simple as inserting a new hard disk drive into the caddy and then mounting it in the system. Other systems require you to purchase a specifically designed drive complete with the proper connections built into it. Replacing the hard drive can be much easier in many portable systems than in their desktop counterparts. This makes it possible for multiple users to share a single machine by simply snapping in their own hard drives. However, because laptops are specialized equipment, any servicing beyond batteries, hard drives, and memory is usually left to specialists or the manufacturer.

The support provided by the system's BIOS determines the upgradability of a system. Older systems, particularly those manufactured before 1995, might offer only limited drive-size options. BIOS chips made before EIDE hard disk drives became the standard can support a maximum hard drive size of 528 MB. A flash BIOS upgrade might be available for your system to provide additional drives. Another option for expanding hard drive space is the PC Card hard drive. This device fits into a Type III PC Card slot and can provide as much as 450 MB of additional space. External drives are also available and can be connected using a PC Card SCSI host or specialized parallel port drive interfaces—you can use any size SCSI drive you choose without being limited by your system's BIOS.

Removable Media

Portable systems are now equipped with other types of storage media that can provide access to large amounts of data. CD-ROM and Zip drives are now available, as well as standard floppy disk drives. Just as in their desktop counterparts, CD-ROM is becoming standard on portables.

The swappable drive bay is increasing in popularity. This product allows the user to switch one of several types of components in the unit. For example, you might not need a floppy disk drive when traveling, so you can insert an extra battery.

Keyboards

Portable keyboards are integrated into the one-piece unit and are therefore very difficult to repair or replace. Unfortunately, the keypad is almost always the first component to fail in a portable. The functionality and durability of the keyboard should be an important concern when purchasing a portable system.

Today's portable keyboards are approaching the size and usability of desktop systems, thanks to the larger screens found in most systems. This has created more space for manufacturers to utilize in the overall design.

Pointing Devices

Today's portable computers come with built-in pointing devices. Most of these pointing devices conform to one of three types: trackball, trackpoint, or trackpad.

Trackball

This small ball (approximately .5 inch in diameter) is partially embedded in the keyboard below the spacebar. The ball is manipulated by the user's finger. These are accurate and serviceable, but they are unpopular because of their tendency to gather dirt and dust, which dramatically reduces performance.

Trackpoint

The trackpoint was developed by IBM and many manufacturers install it in their systems. It is a small, rubberized button (approximately .25 inch in diameter) located above B and below G and H on the keyboard. The user nudges it in any direction (rather like a tiny version of a joystick) to move the cursor around the screen. It is convenient because the user's hands don't need to leave the keyboard to manipulate the trackpoint.

Trackpad

The trackpad is the most recent development of the three—it is an electromagnetically sensitive pad measuring about 1 by 2 inches located in the keyboard below the spacebar. It responds to the movement of a finger across its surface to move the cursor. Mouse clicks are simulated by tapping the pad (buttons are also provided). It's a truly innovative device, but does tend to be overly sensitive to accidental touches and taps. It is also sensitive to humidity, so moist fingers can cause unpredictable performance.

Batteries

A great deal of technology has been developed to extend battery life and improve power management in portable systems. However, battery life is still one of the biggest complaints about portable systems. Even though power management and batteries themselves have improved dramatically over the last few years, the power needed to run faster processors and external devices has increased, leaving battery life about the same. Actual battery life depends as much on how the computer is used, as it does on power-management technology. Simply put, the more you ask the computer to do, the shorter the battery life. Today, battery life is still an issue with portable-system users. Most systems use one of three types of batteries.

Nickel Cadmium (NiCad) Batteries

The oldest of the three technologies, nickel cadmium is rarely used today. It has a shorter life and is sensitive to improper charging and discharging. After being charged, NiCad batteries hold the charge very well. However, their life can be severely shortened if they are not fully discharged before recharging, or if they are overcharged.

Nickel Metal Hydride (NiMH) Batteries

NiMH batteries have a longer life than NiCad (about 50 percent longer), and are less sensitive to improper charging and discharging. They are more expensive than NiCad and don't hold a charge as well when not used. They usually cannot be recharged as many times. They are, however, used in most portable systems, especially those at the lower end of the market.

Lithium Ion (Li-Ion) Batteries

Li-Ion batteries cannot be overcharged, hold a charge well when not in use, and are longer lived than the other two types of batteries. They are also proficient at handling the heavy-duty power requirements of today's higher-end portables. Unfortunately, Li-Ion batteries can be used only in systems specifically designed for them.

CAUTION

Never install a Li-Ion battery in a system designed for a NiCad or NiMH battery. Doing so could result in a fire.

Because they are the most expensive of the three battery technologies, Li-Ion batteries are usually found only in high-end systems.

Sometimes, buying a system with a Li-Ion battery does not mean you will realize a longer battery life. Some manufacturers take the opportunity to make the battery smaller because it is more powerful, thereby saving some space inside the computer while delivering the same performance as a NiCad or NiMH.

New Technology

Battery technology has trailed behind nearly all the other advancements of the portable system. A battery life of two hours is considered very good even when a system's power-saving features are utilized. Some manufacturers are designing systems that hold two batteries to try to overcome this limitation.

A fourth type of battery technology—the Lithium Polymer—has been in development for several years, but it has not yet appeared on the market. Lithium Polymer batteries can be formed into thin, flat sheets and installed behind the LCD panel. They provide approximately 40 percent more battery life while adding far less weight to the system.

TIP

All battery types function best if they are completely discharged before recharging. Even Li-Ion batteries perform better and last longer if they are discharged before being recharged. Another tip is to store charged batteries in the refrigerator. This helps them maintain their charges longer.

Power Management

Some components in a computer system do not need to run continuously. The purpose of power management is to conserve battery life by shutting down these components when they're not needed.

Most portable computers include power-saver modes that suspend system operations when the computers are not in use. Different manufacturers have different names for their power-saver modes such as: *suspend*, *hibernate*, or *conserve*, but they all usually refer to two different states of power conservation: one state continues to power the system's RAM, while the other does not.

Generally, the suspend mode virtually shuts down the entire system after a certain period of inactivity. However, power continues to be supplied to RAM, and the system can be reawakened almost immediately.

The hibernate mode writes the entire contents of memory into a special swap file and then shuts down the system. When reactivated, the file is read back to memory. The hibernate mode takes a little longer to reactivate than the suspend mode, but conserves more battery life. In some systems, the swap file used for the hibernate mode is located in a special partition of the hard drive. If it is inadvertently destroyed, it might require a special utility from the manufacturer to re-create it.

A document jointly developed by Intel and Microsoft—known as the Advanced Power Management (APM) standard—has been, for the most part, responsible for defining the interface (interaction) between the power-management policy driver and the operating system. This interface is usually implemented in the system BIOS.

Another standard currently under development by Intel, Microsoft, and Toshiba is called the Advanced Configuration and Power Interface (ACPI). This standard is designed to place the power-management functions under the control of the operating system. As power-management techniques develop, it becomes difficult for the BIOS to maintain the complex information states needed to run the more advanced functions. Placing power management under the control of the operating system allows applications to interact with the operating system to let it know which of its activities are crucial and which can wait until the next time the hard disk drive is activated.

Chapter 19

Maintaining Computer Hardware

OBJECTIVES

Preventive Maintenance

Safety & Environment

Preventive Maintenance

The best way to repair a failure is to prevent it from happening in the first place. It is a well-known fact of life that failures occur at the most inopportune times. A little preventive maintenance can minimize the impact by either catching and repairing the failure before it happens, or preventing it altogether. The concepts of preventive maintenance and safety have been emphasized throughout this book. Likewise, safety issues are prevalently featured throughout the A+ certification exam. In fact, one objective (10 percent) of the core exam is devoted to this subject. Because of their importance, this chapter focuses on reviewing and expanding upon these issues.

Computers are, by design, very rugged and dependable pieces of equipment. However, like other machines, they age. Several basic procedures, when performed on a regular basis, can prevent premature failures. This lesson discusses how to keep computer hardware running smoothly and reliably.

Cleaning

For the most part, computer equipment is very reliable and lasts a long time. However, as with any piece of equipment, dirt and other airborne contaminants will greatly accelerate the deterioration caused by normal use. Therefore, the best preventive maintenance is to keep the equipment clean.

The first step is to be sure that the computer is installed in a computer-friendly environment. This means that it should be in a dust-free (relatively speaking), smoke-free, and humidity-controlled (within a range of 50 to 70 percent relative humidity) location. For the most part, a normal office environment will qualify as computer-friendly. However, a normal office environment is not the only place that we find computers. Many computers are located on a warehouse floor, in a shop, or grouped together with a large industrial piece of equipment. In the event that the location of a computer is not as desirable as it should be, the frequency of preventive maintenance (cleaning) should be accelerated. In these instances, consideration should be given to establishing a computer-friendly zone around the computer, for instance, installing it into a cabinet and providing a source of clean fresh air. The following table describes what a computer technician should include in a basic cleaning kit.

Item	Usage
Lint-free chamois cloth	A cloth is useful for cleaning the outside surfaces.

Cleaning solution	Simple soap and water should be followed by a clear water rinse. Standard household cleaning solution (not extra strength) can also be used. The solution should be applied to the lint-free cloth and then applied to the computer surface. Do not use aerosol sprays. These generally use solvents as a propellant. Solvents can damage the plastic as well as the electrical components of a computer.
Foam swabs	Use these with cleaning solutions to clean small parts such as the wheels inside a mouse. (Cotton swabs are not recommended, because the cotton fibers can come off and be a contaminant themselves.)
Antistatic spray	An antistatic spray or solution should follow any cleaning. A solution composed of 10 parts water to 1 part common household fabric softener will do.
Small paintbrush and/or small hand-held vacuum cleaner	Used to remove dust from around the computer and inside its cabinet. The vacuum can be used to remove dust from the keyboard and other input devices.
Can of air	For removing dust from the power supply fan or from inside a computer. These can be purchased from any computer supplier; they are made especially for removing dust from electronic equipment.

CAUTION

Never use liquids to clean inside a computer. Never apply liquids directly to the surface of a computer. Never use solvent-based cleanser or aerosols.

The proper placement or location of a computer relative to its environment is important for ease of maintenance and long life. Let's sum up good practices for placement of computer equipment. A computer should be:

- Located in a dust-free and smoke-free environment.
- Subjected to controlled humidity (50 to 70 percent relative humidity).
- Subjected to controlled temperature (do not place too close to a heater or in direct sunlight—avoid temperature extremes).
- Have good ventilation (make sure fan/ventilation vents aren't blocked).

Preventive Maintenance

For the most part, the MTBF (mean time between failures) of a computer and its peripheral devices is quite long. By following the general cleaning and safety measures just described, you can extend this time. This section describes several components and their special maintenance requirements.

Monitors

Monitors require very little maintenance. To keep a monitor in peak condition:

- Keep it clean—use periodic cleaning, dusting, and good common sense with a monitor.
- Use simple cleaning solutions, not aerosol sprays, solvents, or commercial cleansers. DON'T use windows sprays on a monitor screen.
- Do not leave monitors on unattended for extended periods of time. Use a screen saver or the computer's power-conservation features to prevent burn-in of the monitor screen.
- Don't attempt to work inside the cabinet unless you are properly tranied to do so.
- Don't tamper with the monitor. Monitors emit x-ray radiation. Changing the settings or operating the monitor with the cover removed can disable manufacturer's safety devices, thus increasing the hazard.

Hard Disk Drives

Hard disk drives are another type of device that requires very little intervention to keep running. Mechanical failure of hard drives is rare, and when it does occur the solution is generally replacement. The most common problem with hard drives is corrupted sectors. Often they can be repaired with tools such as ScanDisk (part of the Windows 95 and 98 system tools) or one of the many after-market utility software packages available. Here are a few suggestions for preventing problems with hard drives:

- Avoid rough handling.
- Never move a hard disk when it is still spinning.
- Never expose the internal housing to open air.
- Perform regular data backups.
- Use software utilities to maintain the condition of the device (CHKDSK and ScanDisk; hard drive defragmentation programs and antivirus programs).

Floppy Disk Drives

Floppy disk drives are highly susceptible to failure. This is due mostly to the fact that they are exposed to the environment (through the disk slot) and are prone to mechanical damage from inserting and removing the disk. When they fail, the best solution is usually

to replace them because they are inexpensive and simple to install. Here are a few tips to increase the life of floppy drives and disks:

- Do not expose the disks to magnets.
- Never touch the exposed surface of a floppy disk.
- Do not allow smoking near a computer.

Clean the read/write heads. Special head-cleaning diskettes and solutions such as isopropyl alcohol and methanol that do not leave a residue when they dry are available. Cotton swabs are not recommended because of the fibers they shed. Use cellular foam swabs or a lint-free cloth.

Keyboards and Pointing Devices

Keeping a keyboard and mouse clean is key to prolonging their lives. Never place drinks (coffee, soda, tea, and so on) around a keyboard; spilling liquids is a common cause of keyboard failures. Here are a few tips to increase the life of a keyboard and mouse:

- Use a hand-held vacuum cleaner to remove dust from the small crevasses.
- Never use spray cleaners.
- Clean a mouse or trackball by removing the ball and cleaning the *x* and *y* rollers.
- When using a light pen, never touch the ends with your finger.

Printers

Printers are more mechanical than other peripherals and therefore require more attention. Because they use paper, ink, or carbon, printers generate pollutants that can build up and cause problems. Always check the manufacturer's recommendations for cleaning. Here are a few steps for cleaning the most popular types of printers:

Dot-Matrix Printers

- Adjust the print-head spacing.
- Check the tension on the print-head positioning belt. Use a non-fibrous swab dipped in alcohol to clean the print head.
- Clean the printer's roller surfaces.
- Clean the surface of the platen.
- Clean the gear train of the paper-handling motor.
- Apply light oil to the gears using a foam swab.
- Turn the platen to distribute the oil.
- Apply a light coating of oil to the rails.
- Move the carriage assembly to distribute the oil.

Ink-Jet Printers

- Adjust the print-head spacing.
- Check the tension on the print-head-positioning belt.
- Clean the printer and its mechanism.
- Clean the printer's roller surfaces.
- Clean the surface of the platen.
- Clean the surface of the ink-jet print head.
- Clean the gear train of the paper-handling motor.
- Apply light oil to the gears using a foam swab.
- Turn the platen to distribute the oil.
- Apply a light coating of oil to the rails.
- Move the carriage assembly to distribute the oil.

Laser Printers

- Vacuum to remove dust buildup and excess toner from the interior. Remove the toner cartridge before vacuuming.
- Clean the laser printer's rollers using a damp cloth or denatured alcohol.
- Clean the gear train of the paper-handling motor using a foam swab.
- Apply light oil to the gears using a foam swab.
- Distribute the oil throughout the gear train.
- Clean the writing mechanism thoroughly using compressed air. If possible, wipe the laser lens with lint-free wipes to remove fingerprints and stains.
- Clean the corona wires using a swab dipped in alcohol. Be careful not to break any of the strands because if you do, your printer will be rendered useless until they are repaired!

Preventive Maintenance Schedule

There are no universal preventive maintenance schedules that work on every computer. Each schedule must be individualized to meet the needs of the work environment. Use the following suggestions as maintenance guidelines:

Do This Daily

- Back up data.
- Check computer ventilation to ensure that it is clear. Remove any paper, books, or boxes that might impede the flow of air into or out of the computer.

Do This Weekly

- Clean the outside of the case.
- Clean the screen.
- Run CHKDSK or ScanDisk on all hard disk drives. Windows 95 and 98 come with scheduling programs to help you accomplish this on a regular basis.
- Run a current antivirus program and check all drives. These programs also come with scheduling features so this can be accomplished on a regular basis. They will also remind you when to update the virus list (usually done through the manufacturer's Web site).
- Inspect all peripheral devices.

Do This Monthly

- Clean the inside of the system.
- Clean the inside of any printers.
- Vacuum the keyboard.
- Clean the mouse ball and *x* and *y* wheels.
- Defragment all hard disk drives.
- Delete any unnecessary temporary files.

Do This Every Six Months

- Perform an extensive preventive maintenance check.
- Apply an antistatic solution to the entire computer.
- Check and reseat all cables.
- Run the printer's self-test programs.

Do This Annually

- Reformat the hard disk drive and reinstall all software. Don't forget to back up data first.
- Check all floppy disk drives.
- Consider an upgrade to your computer. Check to see that your components can handle your workload.

Safety and the Environment

When maintaining or servicing a computer, several guidelines should be followed to protect you and the environment. This lesson summarizes how to maintain a safe workplace and minimize negative impacts on the environment.

General Safety

Computers and their peripheral devices are electronic equipment, consequently, most safety issues relate to electrical power. However, when you work on this equipment, there are several other concerns to take into consideration, as listed in the following table.

Problem	Prevention
Back injuries	Some equipment, such as printers, monitors, and even the computer itself, can weigh several pounds (10 to 20 pounds or more for newer, larger monitors). This might not seem like much; however, when the equipment is improperly picked up (or dropped), back or other injuries can result. Be especially careful when removing a component from its original packaging. These components are generally packaged very tightly to provide protection during transport and can be difficult to remove.
Cuts	Be very careful when removing covers from computer components. The frames of the cases are often made of thin metal with sharp edges. Also, poorly cut or stamped parts might still have metal burrs, which are very sharp. Devices such as scanners and monitors have glass components that can break.
Tripping hazards	Computers tend to have many cables and wires. If not properly installed, these wires and cables can constitute a serious tripping hazard. Use cable ties to bundle up cables and reduce the "spaghetti" effect. Also avoid running cables under carpets and where people walk.

When installing or working on any equipment, make sure that the work done conforms to all applicable local and national safety codes, such as OSHA (Occupational Safety and Health Administration) and NEC (National Electric Code) standards. Most companies have their own internal safety departments and safety manuals. Be sure that you are familiar with them as well.

Power and Safety

Power is the primary safety hazard encountered when servicing a computer. Be familiar with the following guidelines when working with electrical devices and components.

ESD

The primary electrical-power concern when working with computers is ESD (electrostatic discharge). This subject was covered fully in Chapter 13, "The Basics of Electrical Energy." Remember that while ESD can destroy sensitive computer parts even when the discharge is imperceptible and harmless to humans. If proper ESD tools are not available, touching the case (specifically, the power supply) while working on the computer or its components will provide some protection. However, this will only work if the power supply is plugged into a properly grounded electrical outlet. For a review of power supplies and how to work with them, see Chapter 5 "Supplying Power to a Computer."

Grounds

When used to refer to electronic equipment, the term "ground" can be confusing. Generally speaking, a ground is any point from which electrical measurements can be made. In most cases, a ground means *earth ground*. With early electrical systems, the earth was used as a path for electrical current to return to its source. This is why telegraphs required only one wire (the earth ground serves as the other conductor). In most instances, the frame of the computer is at ground potential or earth ground, as long as the power cord is installed and connected to a properly grounded system. Some electronic equipment uses a special path or conductor for its ground. This is known as signal ground and is not the same as earth ground.

Electronic equipment is both susceptible to and a source of electromagnetic interference (EMI). A properly grounded computer will both prevent the transmission of EMI and protect itself from other sources of EMI. Unchecked, EMI will distort images on a video display, as well as corrupt communications equipment and data on floppy disks.

High Voltages

For the most part, a computer uses ±5 and ±12 volts DC. However, two devices use much higher voltages: power supplies and monitors. With these two exceptions, there are generally no electrical hazards inside a computer.

Power Supplies

The power supply uses 120 volts AC. This voltage is found inside the power-supply case. In most cases, there is no need to open the power-supply case and work on the power supply. The cost of a new power supply is low enough that it is generally easier to replace than repair. However, should you decide to open the case, be careful. Remember, the power switch on most computers (usually located on the front of the computer) also uses 110 volts AC to turn the power supply on or off. If you are working on a computer and leave it plugged in to provide proper grounding, this could present a hazard.

Monitors

Monitors use very high voltages (30,000 volts) to drive the CRT. Remember that monitors are dangerous even when unplugged. They can store this high voltage and discharge it if you touch the wrong parts. Working inside the monitor case should be left to a properly trained technician with the necessary tools.

Power Safety Guidelines

The following are some general guidelines to follow when working around computers:

- Never wear jewelry or other metal objects when working on a computer. These items pose an electrical threat that can cause shorts, which will destroy components.
- To avoid spills, never use liquids around electrical equipment.
- Do not defeat the safety feature of the three-prong power plugs by using two-prong adapters.
- Replace any worn or damaged power cords immediately.
- Never allow anything to rest on a power cord.
- Avoid using extension cords. These can become tripping hazards. Also, they may not be rated to carry the current requirements of the system.
- Keep all electrical covers intact.
- Make sure all vents are clear and have ample free-air space to allow heat to escape.
- Some peripheral devices such as laser printers and scanners use high voltages. Before removing any covers or working on any of these devices, be sure to read the manufacturers' manuals carefully.

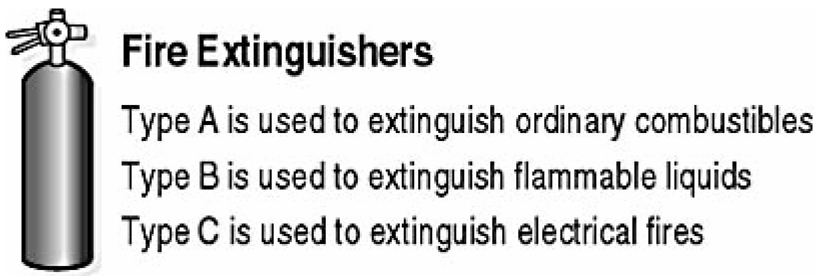
Fire

Fire is not pleasant to think about, but it is a fact of life. A workplace fire can be disastrous both in terms of lost equipment and injury to people. Knowing what to do in the event of a fire can save valuable equipment and, most importantly, lives. Here are a few tips to help prevent fire and protect yourself:

- Always know the emergency procedures to be carried out in case of fire at your workplace.
- Know the location of the nearest fire exits.
- Know the location of the nearest fire extinguishers and how to use them.
- Don't overload electrical outlets.

Simply knowing the location of a fire extinguisher is of no value unless you know how to use it. If you don't, contact your safety department or local fire department. They will be glad to help you get the training you need. Also, remember that using the wrong type of fire extinguisher can be worse than not using one at all.

There are three basic types of fire extinguishers as shown in the picture below.



Fire extinguisher types

Environmental Issues

Many computers and peripheral devices (especially printers) use consumable or recyclable components. In order to help keep our environment safe, you should be aware of these items and use them properly.

Examples of recyclable items or items that require special disposal are:

- Batteries.
- Toner and cartridge kits.
- Circuit boards.
- Chemical solvents.
- Monitors (CRTs).

Be sure to follow the manufacturers' recommendations for recycling or disposal of any of these items. Some items, such as toner cartridges even have prepaid shipping labels so that they can be returned for proper disposal.

When purchasing or using any kind of chemicals (cleaners, for example) that you are not familiar with the proper use and disposal of, be sure to check the MSDS (material safety data sheet). This is a form that describes the nature of any chemicals manufactured. It

includes generic information about the product's chemical makeup and any recognized hazards (including what to do and who to call if there is a problem). These forms are required by law, so ask to see them. Chemical suppliers must provide the purchaser with the MSDS for products, if requested. Also consider purchasing sprays with a manual pump dispenser or compressed air rather than CFCs or other propellants that can be harmful to the environment.

Curriculum Department
Punjab Vocational Training Council
98-B3 Gulberg III, Lahore. Tel: 5872044-48